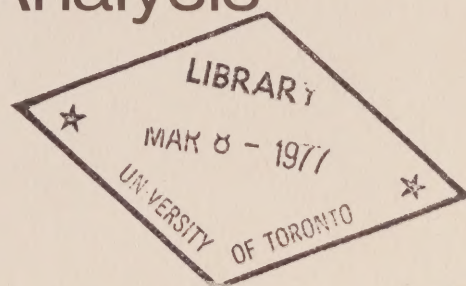


CA24N
DG
.76P67

Government
Publications

Programs and Student Achievement at the Secondary-Postsecondary Interface:

Interproject Analysis



By

H.H.RUSSELL, CARLA WOLFE, PETER EVANS,
RICHARD WOLFE, ROSS TRAUB, ALAN KING

This research project was funded under contract by the Ministry of Education,
and the Ministry of Colleges and Universities, Ontario.

PROGRAMS AND STUDENT ACHIEVEMENT AT THE
SECONDARY-POSTSECONDARY INTERFACE:
INTERPROJECT ANALYSIS


by

H.H. Russell, Carla Wolfe, Peter Evans, Richard Wolfe,
Ross Traub and Alan King

with the assistance of

Marion Marsh

This research was funded under contract by the Ministry of
Education and the Ministry of Colleges and Universities,
Ontario.



Digitized by the Internet Archive
in 2024 with funding from
University of Toronto

<https://archive.org/details/39202312060350>

ACKNOWLEDGEMENTS

The preparation of this report would have been impossible without the dedication and hard work of many people. The authors would like to express their gratitude to the following in particular:

The Ministry of Education and the Ministry of Colleges and Universities, whose comments during the preparation of the report were unfailingly supportive and useful.

The staffs of Project II and Project III and all those who assisted them; without their work, there would have been nothing for us to write about.

Stan Voyce, who coordinated the input and the final printing of the report on the computer, and performed at least a miracle a day in rescuing accidentally deleted computer files.

Irene Leps, who not only administered the project skilfully and efficiently, but went far beyond her duties as Administrative Officer in her contribution of time and work.

Joanne Murray, who assumed the chief responsibility for the final typing of the report, and put in more hours of overtime than of regular time.

Eric McKague and Linda van Will, who took on the job of managing the project during the critical final stages, and brought order out of chaos.

Marie-Noëlle Maillard, whose swift and precise translation was essential to the completion of the chapter on Français and the summary chapter.

Barbara Stone, who produced mountains of typing as the work progressed.

Sheila Scanlan and Elizabeth McLeod, whose assistance with the final typing was substantial, both in and out of regular working hours.

Colleen Voyce, Pam Proctor, Roy Raghunan and Peter Laengert, without whose help the final printing of the report could not have been completed on time.

John Sapey, who kept the computer terminal equipment running with hard work, Scotch tape and prayers.

Wanda Meihm, who walked miles with messages, computer printout, sandwiches and coffee, and saved us many hours.

All the people who regularly use the OISE computer facilities, many of whom delayed their work to allow the Interface staff to make full use of the equipment.

Our wives, husbands and children, whose tolerance and understanding, especially over the last few weeks, were remarkable.

H.H. Russell

Carla Wolfe

Peter Evans

Richard Wolfe

Ross Traub

Alan King

TABLE OF CONTENTS

SUMMARY.....	v
1. INTRODUCTION.....	v
2. STUDY PLAN.....	vii
3. THE DATA.....	x
4. CONCLUSION.....	xxv
REFERENCES.....	xxix
 CHAPTER ONE: INTRODUCTION.....	 1
1. INTRODUCTION.....	1
REFERENCES.....	7
 CHAPTER TWO: STUDY OVERVIEW.....	 9
1. REVIEW OF RELATED STUDIES.....	9
2. RELATIONSHIPS OF PROGRAMS TO TESTS.....	15
3. TOPIC SCORES <u>VS</u> AGGREGATED SCORES.....	16
4. STUDY PLAN.....	18
5. POPULATIONS AND SAMPLES.....	23
6. METHOD OF ANALYSIS.....	26
7. GAPS AND DUPLICATIONS.....	27
REFERENCES.....	31
TABLE 2.1.....	33

CHAPTER THREE: ENGLISH.....	35
1. INTRODUCTION.....	35
2. TEST SELECTION AND COURSE OBJECTIVES...	38
3. PROGRAM EMPHASIS.....	47
4. TEACHER EXPECTATION AND STUDENT PERFORMANCE: TREATMENT OF DATA.....	56
5. READING COMPREHENSION.....	59
6. LANGUAGE ACHIEVEMENT.....	64
7. WRITING.....	73
8. CONCLUSION.....	89
TABLES 3.1 TO 3.13.....	93
APPENDICES.....	125
 CHAPTER FOUR: MATHEMATICS: THE SECONDARY SCHOOL- UNIVERSITY INTERFACE.....	 131
1. INTRODUCTION.....	131
2. ORGANIZATION OF THE DATA.....	136
3. SUMMARY DATA.....	139
4. GAPS.....	147
5. DUPLICATIONS.....	153
6. CONFORMITY OF PROJECT II DATA WITH PROJECT III DATA.....	156
7. SUMMARY.....	157
TABLES 4.1 TO 4.15.....	161
APPENDICES.....	186
 CHAPTER FIVE: MATHEMATICS: THE SECONDARY SCHOOL-CAAT INTERFACE.....	 193
1. INTRODUCTION.....	193
2. PROJECT II DATA.....	195
3. PROJECT III DATA.....	197

4. SUMMARY DATA.....	199
5. INDIVIDUAL TOPICS.....	201
6. DUPLICATIONS.....	207
7. GAPS.....	208
8. SUMMARY.....	215
TABLES 5.1 TO 5.2.....	219
 CHAPTER SIX: PHYSICS.....	 223
1. INTRODUCTION.....	223
2. AN OVERVIEW OF THE DATA.....	227
3. INDIVIDUAL TOPICS.....	229
4. GAPS.....	243
5. DUPLICATIONS.....	248
6. SUMMARY.....	249
TABLES 6.1 TO 6.6.....	252
 CHAPTER SEVEN: FRENCH AS A SECOND LANGUAGE.....	 277
1. INTRODUCTION.....	277
2. PROJECT III DATA.....	280
3. PROJECT II DATA.....	281
4. READING.....	283
5. LISTENING.....	289
6. WRITING.....	295
7. SPEAKING.....	298
8. GRAMMAR.....	306
9. SUMMARY.....	308
TABLES 7.1 TO 7.15.....	310

CHAPITRE HUIT: FRANCAIS.....	331
1. INTRODUCTION.....	331
2. LES DONNEES.....	332
3. PROBLEMES D'INTERPRETATION DES DONNEES.....	334
4. FORME DE L'ANALYSE.....	336
5. <u>LE TEST DE COMPREHENSION EN LECTURE ET</u> <u>DE CONNAISSANCE DE LA LANGUE.....</u>	337
6. <u>LE TEST DE COMPOSITION ECRITE.....</u>	347
7. SOMMAIRE.....	350
TABLEAUX 8.1 A 8.7.....	354
 CHAPTER NINE: ANGLAIS.....	 373
1. INTRODUCTION.....	373
2. THE DATA.....	375
3. READING COMPREHENSION.....	377
4. WRITING EXERCISE.....	381
5. SUMMARY.....	385
TABLES 9.1 TO 9.4.....	386

SUMMARY

1. INTRODUCTION

The purpose of Interproject Analysis is to combine the data from Project II on the nature of students and from Project III on the nature of programs in an effort to identify gaps, duplications or other anomalies which may hamper the effectiveness of the educational system, and to attempt to discover how these anomalies, if any, are related to general public anxiety and to various features of the Ontario system.

Public criticism of both secondary and postsecondary education is not new; in fact, serious criticism of all levels of education has been a continuing phenomenon over a period of decades and even centuries. Current criticism seems to be especially acute in Ontario just as it is in the United States and Europe.

The combination of data from Projects II and III is intended to clarify the ways in which the program characteristics of secondary and postsecondary levels are related to both student performance and the expectations of student performance that are held by instructors from the two levels. However reasonable this sounds, the complexity of the task, when it is attempted on a large scale, is of monstrous proportions. Consequently, the present group of Interface Studies is in many respects a pioneering effort.

The one educational researcher who has provided the greatest amount of information about Ontario's secondary and postsecondary programs and students is W. G. Fleming. His seven-volume series

Ontario's Educative Society and his review of highlights in Education: Ontario's Preoccupation speak to most of the issues of current interest to the Interface Studies. His discussion of the competitive climate in Ontario society through to the 1930s and beyond interacts with his discussion of the traditionalist and progressive philosophical positions of leading educators to show the kinds of balancing which were in effect during successive periods of time.

Two of Fleming's themes are of special significance to the present study. One is the apparent swing toward progressivism as evidenced by the abolition of province-wide examinations, some decentralization of curriculum, and the introduction of the credit system into Ontario secondary schools. The second is the change in composition of the student population in secondary schools and in universities and colleges which has resulted from rapid expansion of the services offered to teenagers.

Because of some highly visible opposition to the apparently progressive credit system, the Ministry of Education commissioned the HSI studies in 1973. In summarizing the findings of those studies, Fleming reported widespread professional and public endorsement of the central idea of flexible course selection. He also found that stability of course and program selection among students was maintained in spite of their freedom to make dramatic changes if they so wished. Fleming and his colleagues also discovered considerable stability in teaching methods, course content and within-school organization. Whether or not there are causal relationships between the swing toward flexible programs and the phenomenal expansion of educational services, it is reasonable to suggest that the latter factor alone stands out as the single most important change of the last quarter-century. Fleming reports a rapid increase in the number of students in secondary schools. The change between the 1945-46 school year and the 1971-72 school year is from 120,000 to 575,000. This is a fivefold increase in gross numbers of students and a change in proportion of secondary school aged students from about 40% to about 80% of all potential students. About half of our current

secondary school students would not have been students at all even fifteen years ago. The schools have been required to cope with increased numbers of the type of student they have traditionally served, and as well they have had to provide service--and perhaps a new kind of service--to as many students again who in the past would not have expected the service at all. There have been profound ramifications to this change, the first and most obvious of which has been the cost to the community, and the second the change in what can reasonably be expected of the average student.

The dramatic changes in numbers and proportions evidenced in the secondary schools are even more pronounced at the postsecondary level. Again according to Fleming, the number of postsecondary students attending Colleges of Applied Arts and Technology and universities stands at about 160,000 in recent years, whereas it was only 19,000 in 1952-53. This is an eightfold increase in numbers, and its effect on the composition of the student body is not clearly known. When the general public, or secondary or postsecondary teachers, fail to take into account the rapid and substantial changes in the composition of the student population, there will indeed be expectations of students which are unrealistic. The data which form the basis for this group of studies suggest that this is precisely what has happened in Ontario, and data reported in the studies mentioned in Chapter Two indicate that other national jurisdictions are encountering very similar problems.

2. STUDY PLAN

There are two basic panels of data for the Interproject Analysis. One pertains to student performance at the secondary-postsecondary interface in English, Mathematics, Physics, French, Français and Anglais; the other pertains to a description of programs at the interface in these same subject areas. The data were gathered according to a plan aimed at assuring a sample highly representative of the entire province of Ontario. With few

exceptions, all Grade Twelve and Grade Thirteen students in Ontario had a chance to be selected for testing in connection with this study; thus virtually all of these students are represented in the aggregated provincial data. The sampling plan which yielded the representative group of students depended on the selection of a sample of schools, and it was this same sample of schools that was used for the selection of the sample of secondary school programs to be studied in the Project III (Nature of Programs) study. This kind of double utilization of a sample placed an added burden on the schools which were selected, but on the other hand it has made possible the analysis of both panels of data on a basis that relates student performance to the courses actually taken by the students, rather than to some provincial norm generated by an independent sample of schools.

The tests chosen to measure student performance were considered the best available by a highly select group of experts from each discipline at both secondary and postsecondary levels. For Grade Thirteen Mathematics and Physics, the choices were particular forms of, respectively, the Ontario Mathematics Achievement Test and the Ontario Physics Achievement Test. Since no suitable test was available of the kind that was needed for Grade Twelve Mathematics, a new test was developed. In English, a writing test was developed, and a multiple-choice language achievement test was constructed of sections of forms of the Canadian Scholastic Aptitude Test, the Canadian English Language Achievement Test and the Canadian Test of English Language. The battery of tests used for French was a very slightly modified version of the tests used with a similar population in the International Study of Educational Achievement. In the Francophone schools, the tests of Mathematics and Physics were identical with the ones used in Anglophone schools, except for language; this was also true of the writing test. First language testing in these schools was carried out by means of a test consisting of sections of the Test de français, langue d'enseignement; the test of second language was composed of a reading comprehension section drawn from the Michigan Test of English Language Proficiency, and a newly constructed section testing writing proficiency.

For each of these tests two Test Appraisal Inventories were developed, designed to gather respectively the opinions of secondary and postsecondary teachers about the tests themselves and about their own coverage of the material tested in the various test items and sections. In responding to the inventories, teachers were asked to keep in mind the courses they were teaching or would be teaching to the students being tested. In one sense these inventories are a modification of the "opportunity to learn" index that has emerged recently as a useful variable in large-scale studies of student achievement.

The basic instruments for gathering program description data are called Course Description Questionnaires; again they were designed to gather information from secondary and postsecondary instructors covering the same courses as above. The Course Description Questionnaires, in contrast with the Test Appraisal Inventories, were developed by committees of subject area specialists; the development of these questionnaires by Project III was done independently of the test selection and the development of the Test Appraisal Inventories by Project II. Thus the Project II and Project III data-gathering activities were to yield two panels of data, with the degree of merging of the two panels dependent on the possibilities of mapping test items and groups of items onto topics and vice versa.

The mapping of Project II test items onto Project III topics was reasonably effective in both Mathematics and Physics; in each of these cases the match between the program and the tests was reasonably good. However, this map was more complicated in the area of language, and here the merging of data is in some respects of a very tentative nature. The merging of data in all of the subject areas was done by means of Topic Tables (see Appendix 4C to Chapter Four), each of which included the data from Project III on a particular topic and the data from Project II on test items related to the topic. On the basis of analysis of the completed Topic Tables for the whole range of topics and test items, possible gaps and duplications and other anomalies emerged. The next step was to prepare summary tables

which included the main relevant information from the Topic Tables and provided indications as to whether topics suffered from any of these anomalies.

3. THE DATA

The Project III data of greatest interest to this analysis are the assessments by secondary teachers of the average level of competence of their students on various topics or objectives at the end of Grade Twelve or Grade Thirteen, and the corresponding assessments by postsecondary teachers of the competence of their incoming students. In every subject area examined, there was a consistent and sizeable discrepancy between these assessments, with postsecondary teachers reporting a much lower level of competence than secondary teachers. Logically, one would expect any discrepancy to work in the opposite direction, at least for the Grade Thirteen-university interface, since those students entering university from Grade Thirteen may be assumed to be, on the average, somewhat more capable than those not proceeding to further studies. (This assumption is strongly confirmed by analyses of score distributions for groups of students differentiated according to grade level and future educational plans--see the Project II report.)

The anomaly can be resolved for practical purposes by ranking the estimates of secondary teachers separately from those of postsecondary teachers, or merely introducing a common discrepancy factor. In either case, the data then reveal high consistency, and prove to be useful in the main analyses. It is interesting, though not surprising, that the correspondence between ratings and student performance is closer in the case of secondary teachers than for postsecondary teachers.

Why the discrepancies arise is not so easily resolved. Perhaps it is a lag on the part of university instructors in accommodating to the changes in the composition of the student

body discussed in Chapter Two--changes which are both more pronounced and more recent in universities and community colleges than in secondary schools. Or it may be nothing more than the fact that the two sets of teachers are responding from different contexts--i.e., that knowledge which in contrast with that expected in lower secondary grades appears to be at a high level looks quite different when examined in the context of advanced university work. The Project III staff found some evidence that this latter explanation is at least part of the answer: secondary and postsecondary teachers, presented with the same test questions and asked to rate them according to the level of knowledge they demanded, differed somewhat in those ratings. However, these differences were not so great as the differences evidenced in the topic ratings.

Student performance data from Project II presented one problem in particular--that of deciding what constituted "acceptable" student performance on a test or a test item. The decisions were necessarily highly arbitrary, and varied from subject to subject in accordance with the difficulty of the tests and the level of the students.

Project II also provided Test Appraisal Inventory data for each test. For Mathematics, Physics, and some parts of the language tests, secondary teachers were asked to classify each test item according to whether its content should have been known by students entering their courses, should have been learned by all students in the course, should have been learned by some students in the course, or was not expected to have been learned before or during the course. Postsecondary teachers rated the items as old knowledge, course content, or other. This method of classification could not be easily applied to most parts of the language tests. Therefore, more general questions were asked about the tests, about the amount of emphasis the teachers placed in their teaching on the development of the particular skills required to respond to the test items, and about the proportion of students who should be competent in these skills.

3.1 Mathematics--Grade Thirteen

In the case of Mathematics at the Grade Thirteen-university interface, there must be some caution introduced in making historical comparisons, because of curriculum changes. In spite of this, the evidence is strong that student performance on the test employed has been relatively stable over the eight years since the original administration of the test. In fact, scores for university-bound students were slightly higher than for the 1968 administration. It appears, therefore, that present curriculum problems in Mathematics would not be solved by returning to programs or conditions of the earlier era. It may be comforting to teachers and school authorities to know that the widespread anxiety about declines in student performance in this field is generally unfounded. On the other hand, however, the data indicate a number of problems, the resolution of which could be very helpful to our students.

The test employed in this study is not ideal for measuring student achievement in Ontario Grade Thirteen Mathematics courses. One-quarter of the test items deal with material covered in the curriculum only before the Grade Thirteen level, and it could be argued that treatment of such material is inappropriate in a test at this level. In the field of relations and functions coverage is also given to material treated both before and during Grade Thirteen. However, particularly in view of the number of cases in which such items revealed inadequate student knowledge in the present study, it does in fact seem important to test this sort of material.

A more serious problem with this test is its lack of coverage of a number of topics dealt with in Calculus, and of the entire content of Algebra. These areas should certainly be dealt with in any future standardized testing program. It would also be wise, in view of differential enrolment in these courses, to use separate tests for each course. In this study, for example, it was not possible to test students enrolled in only one of

Calculus and Relations and Functions, since the test material could not be separated.

In Mathematics at this level, a number of gaps were pointed out by items on which the student success rate was low, and for which a sizeable percentage of university teachers expected incoming students to have a knowledge of the content. These gaps were evidenced in the following topics:

- (a) I-1: Function as a mapping (1 item of 2)
- (b) I-2: Inverse of a function (1 item of 1)
- (c) I-3: Graphs and properties of second degree relations using previously known skills (1 item of 4)
- (d) I-8: Intersection of conics and conics (1 item of 2)
- (e) I-10: Standard trigonometric formulae and applications (4 items of 4)
- (f) Pre-Grade Thirteen material (3 items of 9)

Duplications were indicated by items where the success rate was high, and where significant numbers of university teachers stated that they retaught the topic or performed (apparently unnecessary) review on it. Duplications occurred in the following topics:

- (a) I-6: Applications (of equations of conics in non-standard positions) (1 item of 1)
- (b) II-2: Rate of change: slopes, secants, tangents (2 items of 2)

- (c) II-3: Derivatives of powers, products, and quotients
(1 item of 1)
- (d) II-8: Maxima and minima problems (1 item of 1)
- (e) II-28: Derivatives of elementary functions (1 item of 1)
- (f) Pre-Grade Thirteen material (1 item of 9)

It is important that most of the duplications involve calculus items, and most of the gaps are related to topics in relations and functions. It is perhaps even more important that almost all gaps could be traced back to levels before Grade Thirteen. In these cases, the item content had been considered old knowledge on Grade Thirteen entry by most teachers, and had not been widely reviewed or retaught.

Some approaches at the secondary level to achieving more efficient learning in mathematics suggested by this analysis, then, are a greater emphasis at pre-Grade Thirteen levels on mastery of those skills needed in the senior courses; a more careful examination by Grade Thirteen teachers of the extent of students' previous knowledge; and more effort by these teachers to remedy any existing gaps in basic knowledge. At the university level, teachers might safely put more confidence in their students' knowledge of calculus, and use the time saved in this way in more extensive review and teaching of those areas in which student competence appears to be lower.

3.2 Mathematics--Grade Twelve

In general, the performance of students on the test constructed for use in measuring mathematical skills in Grade Twelve supports the usual complaint of CAAT teachers that their incoming students do not possess the degree of skill they should have in doing basic

arithmetic and algebra. In certain cases, teachers at the college level appeared to expect more than a reasonable amount of their incoming students, many of whom would not have studied Mathematics at the Grade Twelve level (although it should be pointed out that Grade Twelve Mathematics is a prerequisite for technical courses at the CAATs, and teachers of such courses are justified in expecting more knowledge of their students). But much more often these students appeared not to have mastered material taught to them in the earlier years of secondary school, when all or almost all of them were still enrolled in Mathematics courses. The situation is more serious than the list of gaps resulting from this Interproject Analysis indicates, since the sample of students included many who planned to continue their secondary studies, presumably with the aim of eventually attending university. Project II data indicate that these are on the average the more able students, and that the performance of those students making the move from Grade Twelve into the CAATs was somewhat below the overall performance reported here (see Project II report).

The distribution of gaps over the general topics tested was:

- (a) I: Basic Arithmetic--total of 6 items; 1 minor gap, 2 major gaps.
- (b) III: Basic Algebra--total of 19 items; 4 minor gaps, 4 major gaps.
- (c) IV: Quadratic Functions and Equations--total of 2 items; 2 major gaps.
- (d) V: Exponential and Logarithmic Functions--total of 4 items; 1 minor gap, 2 major gaps.
- (e) VII: Analytic Geometry and Vectors:--total of 4 items; 1 major gap.

(f) Topic not listed:--total of 1 item, 1 major gap.

The general impression given by the comparison of performance data with data on teacher expectations, even for this larger group containing many students more advanced than those going on to CAATs, is discouraging. There was no topic for which all related items were answered satisfactorily by the students, although the situation approached the satisfactory for a few. In three topics, some or all of the gaps can be attributed to overly high expectations at the CAAT level that students will be familiar with material which is not widely taught before Grade Twelve; our assumption is that a significant number of CAAT entrants will not have continued their mathematical studies to this level (except, as noted above, in the case of technical courses). In most topics, however, the problems can confidently be assessed as the result of inadequate learning in the earlier grades of secondary school. It is disturbing that in many cases the skills have not been given adequate attention in the form of review in Grade Twelve in spite of the clear deficiencies in student knowledge. (As noted above, most of the gaps found at the Grade Thirteen level in Mathematics could also be traced back to previous years, where material was not properly mastered; again the teachers of the courses in the final secondary year assumed that their students knew this material and often did not review it.)

Remedial action could be taken at all three levels. In the earlier secondary years, more rigour is required to make sure that students have some understanding of how to apply the skills they are taught which they will need at higher levels. Grade Twelve teachers should spend more time reviewing these skills and assuring that students can use them competently. CAAT teachers should resign themselves to the fact that, except possibly for those entering technical courses, incoming students will have an imperfect knowledge of exponents and quadratic equations in particular. If the students are known to be competent in the more

basic areas, these teachers will require much less time for review and can use the time saved to cover these two problem areas and to proceed to more difficult work. An alternative, of course, is to require all secondary students planning to enter community colleges to carry their enrolment in mathematics courses to the Grade Twelve level. Presumably, given the other remedial possibilities discussed above, they would then be skilled in exponents and quadratic equations as well, and would provide the CAAT teachers with even more time for new work.

3.3 Physics

In contrast to the situation in Grade Thirteen Mathematics, where performance on the test used remained relatively stable from 1968 to 1976, Physics showed something of a decline. Between 1970, when the test was first used, and the present administration, the scores for Grade Thirteen students planning postsecondary education dropped an average of three to four points out of 60. The same cautions must be used here as in the case of Mathematics in making historical comparisons, but a problem is certainly indicated; its causes are unclear. One possibility is that the decline results from a decrease in the time allotted to Physics classes.

Of the 21 general topics included in the list used in Project III, the standard Grade Thirteen physics course deals to at least a limited extent with 13, frequently omitting a number of individual topics from a section. The Project II test contains items related to 11 of these 13 general topics, omitting only Electricity and Magnetism and Atomic Structure. Its 60 items touch on 38 of the 60 individual topics included in the course under these 11 general headings.

On historical evidence, a somewhat arbitrary scale was established for item results. Using this scale, student performance was very good on five items, good on seven, adequate on 15, poor on 26, and very poor on seven.

These judgements of item results must, of course, be modified by a number of factors. A poor success rate on an item whose content has not been taught to many students means something quite different from a similar result on content taught to everyone. Student performance on an item whose content most university teachers expect them to know is more crucial than on one whose content is not required knowledge on admission to a first year university course. The data were therefore examined with a number of such factors in mind.

It was found that the problem of duplication--of material already known by the student being taught to him/her at the university level as new--was very rare, to judge by the test results. In only three cases was item content widely taught as new at that level when student performance on the item was good or very good.

A substantial number of items, however, suggested the existence of gaps between student knowledge on admission to the university and teacher expectations of the knowledge of incoming students. There appeared to be four different kinds of gaps, scattered over all topics tested. In the first case university teachers expected students to have command of material which was given little emphasis at the lower level in compliance with the Ministry of Education guidelines for the course. In the second, even though these guidelines recommended major emphasis for a particular topic, 20% or more of secondary teachers did not cover the content of the relevant items with all of their students. In the third, Grade Thirteen teachers assumed a much higher level of previous knowledge from their entering students than was the case, if one may judge by the fact that students were often unable to handle test items whose content was described by secondary teachers as "old knowledge". The fourth and largest group of problem items, however, dealt with material widely taught at the Grade Thirteen level, but whose content had not been mastered by students even though the guidelines suggested that the topic receive major emphasis.

For Physics, the distribution of items suggesting gaps was as follows:

- (a) 1: Measurement (5 of 6 items)
- (b) 2: Functions (2 of 3 items)
- (c) 3: Motion (Kinematics) (12 of 18 items)
- (d) 4: Newton's Laws of Motion--Dynamics of a Particle (5 of 7 items)
- (e) 7: Gravity--Near the Earth's Surface (4 of 4 items)
- (f) 8: Universal Gravitation (2 of 2 items)
- (g) 9: Momentum (5 of 8 items)
- (h) 10: Work, Energy and Power (6 of 9 items)
- (i) 11: Vibrations and Waves (7 of 8 items)
- (j) 12: How Light Behaves (4 of 4 items)
- (k) 13: Interference and Diffraction (4 of 8 items)

If the gaps adumbrated by this analysis are real, then it appears that responsibility for them must be divided. Teachers at the earlier levels of the secondary school should be ensuring that students leave their courses with a good command of basic physical principles and relations. Those at the upper level should first of all be checking that their incoming students have this basic knowledge, which appears lacking in many cases, and then reteaching the material where necessary. (This may be of especial importance in Physics, since the usual pattern of study leaves a one-year period between the students' previous Physics course and

the Grade Thirteen course, and a great deal may be forgotten in a year.) All teachers should be covering with the appropriate amount of time and attention those topics which the guidelines state to be worthy of major emphasis. And university teachers should not expect high competence in other areas of the subject, which must of necessity be short-changed if sufficient time is to be spent on the more important areas.

3.4 English

With English as a first language, as with the other language areas studied, it is important to keep in mind the difficulties encountered in relating specific skills (Project II data) to specific topics or objectives (Project III data). The kinds of tests and instruments employed did not lend themselves to this type of correspondence. Also, it was necessary to restrict evaluation of student performance in English to the specifics of reading, language and writing, leaving aside the many objectives related to literature, which evidently compose a high percentage of content in many courses at the interface. Consequently a review of apparent anomalies, gaps or duplications must be severely restricted in relation to all possible course objectives, and must refer to objectives in rather general terms.

In the interproject analysis of English it will be apparent that no significant duplications have emerged. Expectations generally appear either to be in conformity with student performance, or else overly optimistic. There may indeed be duplications in the area of literature as so high a proportion of time and emphasis is given to this area of English, but this is not discernible from an analysis that has had to be confined to reading, language and writing.

One "gap" in testing in the areas of language and reading has been emphasized at different points in the chapter dealing with English, but should probably be underlined here. It is evident that the language test employed in this study--the most

appropriate available--is too difficult for all but the university-bound student. It does serve no doubt as a screen to separate that student from the rest, but it is not helpful in characterizing performance in language or reading below that level. As well, responses to the Test Appraisal Inventories indicated that teachers felt some discomfort with the format of this test, and supported very strongly the inclusion of a sample or samples of student writing in any appraisal of language performance.

The various indices used to analyze gaps at the interface appear to indicate that in most aspects of "literacy" examined, secondary and postsecondary teachers have made a fairly realistic estimate of the competences in English of SSGD and SSHGD students. Though all teachers would like to see performance better than it is, we find fairly strong evidence of a gap in expectation only for reading ability at the Grade Twelve-CAAT interface. It is probably true that first year CAAT teachers have higher expectations concerning the reading ability of their incoming Grade Twelve graduates than are warranted. Some strategies should probably be found--perhaps more time for reading, perhaps a more formal program of reading instruction--to improve the reading skills of students in general level programs. This gap does not appear to extend to SSHGD students proceeding to university; the relatively greater emphasis placed there on inference skills than on reading for literal meaning appears to be in accord with performance.

The main gap discerned was the distance between what teachers in this study say should be emphasized in writing, and what in fact appears to be the case in terms of time given to writing in a high proportion of courses in all but the CAATs.

Serious errors in grammar, word choice and sentence structure did not characterize the writing of students above the 25th percentile at either interface, and students generally showed competence in organizing what they had to say.

The two serious concerns in the samples of writing examined closely were, first, the frequency of errors in conventions, and second, the general dullness of the writing.

One direction that might be followed to help close these gaps is a shift in course emphasis to provide more writing opportunities. As well, something can be said about the kind of emphasis needed in writing programs. There does not appear to be a need for analysis, drills, and reviews of learning about the nature of the sentence. Rather the emphasis should be on the creation of more effective stimuli for writing, and the encouragement of the student to write with commitment for a real audience and to experiment with diction and style. A writer with a strong sense of audience is likely to show more respect for the conventions of written English as well. Lastly, greater emphasis on language and writing in the program would ensure that the students have more frequent opportunities to write.

3.5 French

In the study of French as a second language, there was a generally good fit between teacher expectations and student performance in both reading and listening skills. Performance was consistently at an acceptable level in these areas.

This pattern is also true of pronunciation, but a different pattern emerges when it comes to other skills connected with speaking, and to writing. Although average performance in writing and speaking was at a reasonable level, the amount of variability in performance was very great, ranging from almost total inability to speak or to write fluently to a very high degree of facility in using the language. In these areas one can only say that there exist simultaneously a gap and a duplication. The large number of students performing poorly cannot help but be out of their depth in a university course; the equally large number performing well must be frustrated by the need for university

teachers to devote time to bringing others up to the same standard of performance.

Somewhat the same situation holds with respect to grammar; variability was great here as well. However, the average standard of performance was lower here than would be suggested by the writing and speaking results as a whole, and consequently this is more properly defined as a gap, with some accompanying problem of duplication for those students who have achieved a high level of mastery of morphology and syntax.

In French as a second language, then, the "passive" skills of reading and listening were handled satisfactorily and with relatively little variability; performance largely matched teacher expectations. The "active" skills of writing and speaking were problem areas (except for pronunciation). Average performance on writing and speaking, excluding the component of grammatical correctness, was adequate, but with very high variability. Average command of grammar was not adequate, and the same degree of variability was evidenced.

3.6 Français

In Français (French as a first language), the greatest difference in competence between Grade Twelve and Grade Thirteen students was manifested in vocabulary. There was a marked difference in scores for the two groups on the section of the language test concerned with vocabulary, and on the evidence of the score distributions it must be assumed that this difference resulted largely from very poor performance by Grade Twelve students not intending to pursue secondary education, but rather planning either to begin postsecondary work (presumably at a CAAT) or to leave the system. This assumption is reinforced by a comment of those researchers who engaged in a close examination of some of the writing produced by the students tested; they remark that some students had a very poor command of vocabulary (see Appendix B to the Project II report).

Some weakness was evidenced at both the Grade Twelve and Grade Thirteen levels in grammar. The language test concerned itself specifically with spelling as related to agreement (subject-verb, adjective-noun, etc.). In spelling, both here and in the writing produced by the students, there were apparent difficulties in spelling caused by the fact that words pronounced alike are spelled differently (e.g., "c'est/s'est"). Subject-verb agreement appeared to be a particular problem in grammar, especially in cases where the subject was somewhat complicated in form. Many errors in usage, especially of prepositions or verbs, appeared to be attributable to interference from English. The frequency of errors in the writing test must be considered unacceptably high, with even those Grade Thirteen students planning to attend university producing an average of one error of some kind per ten words.

A general comment made on the style evidenced by the students in the writing exercise is that students showed a strong tendency to use in formal writing many expressions and structures unsuitable to that mode, although acceptable in informal French. This was attributed at least in part to the relative lack of exposure to formal French which is a result of the minority status of Franco-Ontarions (see the Project II report, Appendix B).

The same readers commented that the level of logical presentation of argument in these papers was very low--that in fact only a tiny minority of students showed a capacity to argue a position in an organized and logical way.

Reading was generally well-handled at both grade levels, except for one reading comprehension passage in which many students apparently failed to detect the ironic tone, and consequently gave answers to the related questions which were based on a literal interpretation of the text. Some emphasis is indicated on the ability to appreciate such subtleties of tone.

3.7 Anglais

In Anglais (English as a second language), the students tested performed at what appeared to be a quite acceptable level in reading. Their ability to write a summary of the longer passage presented to them in the second half of the test appeared to be somewhat less than their reading ability, but performance was consistent and not unacceptable. Great variation was manifested, however, in the portion of the writing exercise which required them to state and defend their own positions on the subject discussed in the passage. Grade Twelve performance, particularly among those students probably enrolled in the less demanding courses, was far below the quite acceptable standard set by the Grade Thirteen students.

4. CONCLUSION

One conclusion, which is carefully documented in the case of Grade Thirteen Mathematics, is that there is no golden era of the past, and no greener field far away, offering readymade solutions to the problems encountered by school programs in the basic subjects. There is strong evidence that the group of students passing through the interface between secondary and postsecondary studies is as well-educated and as well-prepared in basic skills as were similar groups in Ontario in the past,¹ and as are comparable groups of students in other countries.²

As we have pointed out, probably the most striking change in education in Ontario over the past 25 years has been the dramatic increase in the numbers and proportions of young people availing themselves of the right to education in the senior years of secondary school and in postsecondary institutions. This phenomenal expansion has placed unprecedented strains upon the province's resources of both money and manpower, and one of the effects of the resultant strain may be the kind of public anxiety now being evidenced.

One of the most important factors in this expansion has been the inclusion in the student population of those groups who are enrolled in the General courses in senior secondary school, and in Colleges of Applied Arts and Technology. If we examine those students in this group who do not appear to have learned very much, or who are greatly dissatisfied with their schooling, we can perhaps discover where the system has room for improvement; in fact, we may be confident that improvements of this sort are almost a necessity if we are to continue to have public support for education. If on the other hand we examine those students in this group who have learned well, and who are satisfied with their education, we have evidence of the kind of success that the expansion of our educational system was designed to achieve. We have many students who would not in the past have been served at all by the system after their early teenage years; now they are served well, and society reaps its reward from them in increased potential for productive and useful work. This second group of students is the profit we have gained from expansion; the first group may presently be counted as our loss, but presents us with the potential for future gain.

It is clear from the data presented in this report and in the Project II report that even the best tests available for use in this study were inadequate to the task of measuring student performance throughout the domain of their studies. If decisions about education are to be founded on data that are amenable to scientific analysis, the development of tests to measure student performance at the end of secondary school is a necessity, and should be given high priority.

Even with these imperfect instruments, however, it is clear that real gaps, duplications and other anomalies were evidenced in virtually all subject areas. It is important for teachers at both sides of the interface to study the data, and to attempt to come to some kind of resolution of the current problems. The identification of particular gaps, duplications and anomalies in this 1976 study may in the long run be of less significance than the concurrent development of the techniques for gathering the

needed data. We are not suggesting that this technology is advanced--rather, it is only emerging. However, even at this stage it is adequate to the task of annual or periodic monitoring of the provincial programs, and is useful as well as a prototype for the monitoring of programs by school systems.

What we have seen in this study is clear evidence, from a number of types of data and from a number of perspectives, that the perceptions of various publics with an interest in education do not entirely correspond with reality. That there are problems in our educational system at the secondary-postsecondary interface is not in dispute. However, many of these problems are not new, and most of those that are of recent origin are direct consequences of the dramatic increase in the population of our schools, colleges and universities. We hope that our data will point out some of the problems, suggest some possible solutions, and form the foundation for fruitful debate on the future direction of education in Ontario.

FOOTNOTES

¹We have already mentioned the stability of results in Grade Thirteen mathematics. Although this sort of hard evidence is not available in other subject areas, it may be of interest to quote from The President's Report 1950-1951, University of Toronto; the words are Sidney Smith's:

The Department of English at University College recently gave a simple examination to students in all three years of the Pass Course who had elected English as one of their subjects. The examination was designed to test the student's knowledge of punctuation, the range of his vocabulary, and

his ability to summarize a piece of expository prose. The results were, in the highest degree, alarming. The failure rate was 65 per cent. Certainly there is no reason to expect that the failure rate would be any lower if a similar examination were given to students in the Faculty of Arts who do not take English, or to students in the professional faculties. Two conclusions present themselves: the training in English given in the high schools is no longer an adequate equipment for work in the university; and it appears to be incumbent on the university to take immediate steps to make sure that illiteracy no longer dwells in easy partnership with the possession of a degree.

²The second conclusion of the International Study of Student Achievement is: "Between countries with long established universal educational systems there is not much variation in either the average or the range of achievement. In all these countries there are both children who have learned a great deal and others who have learned little." (Peaker)

REFERENCES

- Fleming, W.G. Education: Ontario's preoccupation. Toronto: University of Toronto Press, 1972.
- King, Alan J.C., et al. Continuity and diversity of courses. The secondary-postsecondary interface project III: Nature of programs. Kingston: Queen's University, 1976.
- Peaker, Gilbert F. An empirical study of education in twenty-one countries: A technical report, p. 172. Stockholm: Almqvist and Wicksell International, 1975.)
- Traub, R., Wolfe, R., Wolfe, C., Evans, P., Russell, H.H. Secondary-postsecondary interface project II: Nature of students. Toronto: The Ontario Institute for Studies in Education, 1976.

CHAPTER ONE

INTERPROJECT ANALYSIS

1. INTRODUCTION

The nature of programs and the nature of students in Ontario secondary schools, and to a lesser extent colleges and universities, have been examined in previous research (Fleming, King, Laxer, Ryan, Leithwood, Alexander). In spite of the introduction of widely acclaimed changes, the main features of the Ontario system have remained relatively stable over long periods of time, as has also been the case in the rest of Canada and in other western countries (Peaker). Public criticism of both secondary and postsecondary education is not new and, in fact, serious criticism of all levels of education has been a continuing phenomenon over a period of decades, and even centuries. Current criticism in Ontario seems to be especially acute just as it is in the United States (Harnischfagen and Wiley) and Europe (Peaker).

Ontario's Grade Thirteen, which leads to an Honour Graduation Diploma, is a feature which distinguishes this system from most others. It is significant, therefore, in any consideration of the secondary-postsecondary interface. However, the fact that criticism of the secondary schools seems to be as acute in jurisdictions where there is no Grade Thirteen as it is in Ontario suggests that the present problems, if they are in fact problems, are of a more general nature than can be related specifically to this one feature of the Ontario system.

The introduction of Colleges of Applied Arts and Technology during the mid-1960s was a new feature of the Ontario system and thus is worthy of special attention in the Interface studies. Although the Ontario colleges possess some unique features, in many respects they are parallel to college systems in other large jurisdictions, and encounter similar problems. The introduction of the college system is, in fact, partly responsible for a significant change in the proportion of teenagers who remain in school, whether in Grade Twelve, Grade Thirteen, community college or university.

The increasing proportion of young people attending school reflects a trend in western democracies toward providing educational service more equitably, and although there are national differences in the rate of progress toward the goal of full education for all future citizens, the problems created by the trend seem to be common to many countries (Peaker). Briefly, the problem is that the senior grades of high school and the postsecondary educational institutions are providing educational service for not only the academic elite who, by tradition, have expected such educational provision, but also the students of more limited academic capability who traditionally have not expected such service. The result has been an apparent decline in standards of student performance, and the general public may believe that this is a genuine decline which affects their children adversely, when in fact it may well be a real advancement which benefits their sons and daughters by increasing both their array of marketable skills and their general educational maturity.

The purpose of the Interproject Analysis Study is to combine the data from Project II on the nature of students and those from Project III on the nature of programs, in an effort to see whether there are gaps, duplications or other anomalies which affect the system adversely. How these anomalies relate to the general public anxiety or to the various features of the Ontario system may be clarified as the study proceeds but, as the study begins, the range of possible outcomes is extensive.

The contract specifications for Interproject Analysis include the following:

1. The determination of the content validity of the achievement tests administered in Project Two, with particular emphasis on the identification of course material not covered by the tests.
2. The identification of the range of expectations concerning student performance, and the explanation of differences in expectations in terms of differences between schools, programs, courses and topics within courses. (Such factors as teacher's pedagogical background and the socio-economic status of the students might also be considered as factors affecting expectations.)
3. The identification of the nature and extent of gaps and duplications across the Interface resulting from differences in expectations, and an assessment of the impact of these gaps and duplications in terms of the number of students who experience them.
4.
 - (A) The identification of differences between (i) instructors' expectations for appropriate student achievement upon secondary school graduation or upon entrance to a postsecondary program, and (ii) the actual achievement of the students as measured by achievement tests.
 - (B) The explanation of these differences in terms of differences between schools, programs, courses, and topics within courses.
 - (C) The separation of student-related causes from program-related causes of these differences between expectations and achievement.

Areas two and three describe analyses which the Interproject Analysis will be able to achieve over and above that done in Project Three.

1.1 Problem Statement

In what ways are the program characteristics at secondary and postsecondary levels related to student performance and the expectation of student performance that is held by teachers from the two levels? This general problem statement encompasses most of the contract specifications and is useful in relating the plan of the study to both the formal outcome requirements and the informal issues of today.

1.2 General Approach of Study

One of the first problems that the investigators encountered in planning the Interproject Analysis Study was that of determining the amount of detail to be reported in the data on both student performance and program in order to present relatively clear conclusions within the time and manpower limitations. On the basis of the combined experience of the investigators, it was decided to conduct the analysis on an item-by-item and topic-by-topic basis. Although this is a departure from tradition, it was the investigators' view that there would be enough variation among topics and among test items so that some important issues would be clarified that otherwise would be clouded through the use of traditional aggregatory procedures. It appeared feasible to conduct this detailed analysis of subject matter, but in the domain of students and courses the pressure for more detailed analysis seemed only to lead to chaotic sampling and data-gathering problems. In the case of students, the size of the sample would mushroom to unmanageable proportions if all sub-populations were to be adequately sampled. In the case of courses, at the postsecondary level in particular, there seemed to be such a high degree of uniqueness that no amount of

sampling could remove distortions. A general solution to the problem of detail seemed to be to aggregate students and courses across the province while at the same time, where possible, disaggregating tests to the level of the test items, and programs to the level of topics or objectives. One important feature of this decision is that it focusses on variables such as topics or test items for which remediation by educators is possible rather than on those variables such as students, teachers or schools for which remediation is either impossible or not in the interests of the majority of of the population.

The Project II data provide evidence as to the performance levels of secondary school students in a variety of subjects and, within these subjects, on a variety of topics or test items. In order to relate these hard data pertaining to student performance to the judgemental data of Project III, it is necessary to map the test content (Project II) onto the program content (Project III) and vice versa. As a result of our needs for very detailed program and student performance accounting, each test item from the Project II data was mapped onto each topic from Project III data that could be shown to be related to the test item. This mapping made it possible to relate the Project III data, revealing the secondary and postsecondary teacher perceptions of student requirements and program requirements on these topics, back to the student performance data of Project II. The kind of "orderliness" of data that is envisioned to this point materializes more or less as planned in both the mathematics and the physics chapters, but the nature of the content, the nature of the data and the lack of clear mappings leads to somewhat different procedures for analyzing data in the language chapters.

The kind of analysis envisioned thus far involves a detailed breakdown of the subject areas into topics and test items, and at the same time the accumulation of data from a representative sample within the province. Variations across regions and variations across classes within a region could have been studied if a larger sample had been chosen and if the sampling procedures had been focussed on the various data of interest. The present

study does not permit such a breakdown on a representative basis, but a case study of two regions has been carried out and, within the sampling limitations that prevail, some of the variations of secondary interest will be suggested.

REFERENCES

- Alexander, W.E. and Farrell, J.P. The individualized system: Student participation in decision-making. HSI Studies. Toronto: The Ontario Institute for Studies in Education, 1974.
- Fleming, W.G. Education: Ontario's preoccupation. Toronto: University of Toronto Press, 1972.
- Fleming, W.G. The individualized system: Findings from five studies. Toronto: The Ontario Institute for Studies in Education, 1974.
- Harnischfeger, Annegret, and Wiley, David E. Achievement test score decline: Do we need to worry? Chicago: CEMREL Inc., December 1975.
- King, Alan, J.C., et al. Continuity and diversity of courses. The secondary-postsecondary interface project III: Nature of programs. Kingston: Queen's University, 1976.
- Laxer, G., Traub, R.E., and Wayne, K. The individualized system: Student social and achievement patterns. HSI Studies. Toronto: The Ontario Institute for Studies in Education, 1974.
- Leithwood, K.A., Clipsham, J.S., and Davies, C. The individualized system: Courses and patterns of student choice. HSI Studies. Toronto: The Ontario Institute for Studies in Education, 1974.
- Peaker, Gilbert F. An empirical study of education in twenty-one countries: A technical report. Stockholm: Almqvist and Wiksell International, 1975.
- Ryan, Doris W. Administration and Leadership. HSI Studies. Toronto: The Ontario Institute for Studies in Education, 1974.

CHAPTER TWO

STUDY OVERVIEW

1. REVIEW OF RELATED STUDIES

One educational researcher, W.G. Fleming, dominated the field in Ontario during the 1960s. His seven-volume series, Ontario's Educative Society, and his review of highlights in Education: Ontario's Preoccupation provide a detailed account of virtually all of the historical issues that have a bearing on the current secondary-postsecondary interface issues. His discussion of the competitive climate in Ontario society through to the 1930s and beyond interacts with his discussion of the traditionalist and progressive philosophical positions of leading educators to show the kinds of balancing which were in effect during succeeding periods of time. The abolition of province-wide examinations at the end of Grade Eight, and at the end of Grade Thirteen two decades later, are just two indications of the force of progressive ideas.

The historical and philosophical background provided by Fleming is thorough, and some of the themes warrant attention in the present investigation. The swing toward progressivism in theory, in contrast with stability in the practice of teaching and the selection of subjects by students, is one such theme. Another is the rapid expansion of educational service to potential secondary and postsecondary students. The progressive ideas which had found official endorsement in some government policies in the 1930s maintained a little momentum through the forties and fifties, and then achieved even more prominence in the 1960s. The changing composition of the student population in secondary

schools and the changing needs of society helped in the introduction of official policies which gave the Ontario high school of the 1970s a very progressive look.

The rigidity of program content and course selection that characterized the final year of high school up until 1967 was intended to be replaced by a much more flexible plan, partly through the abolition of the province-wide Grade Thirteen examinations. The credit system, which was formally introduced on a province-wide basis in 1972 under the authority of the document HSl, allowed a student to earn a secondary school honour graduation diploma with a fixed number of course credits, with very few limitations on his/her choice of courses. The opponents of the apparently progressive credit system received widespread media coverage of their objections to the program. Some of them claimed that high schools' success in teaching basics declined dramatically almost overnight.

The Ministry of Education commissioned the HSl studies in 1973. In summarizing the findings of those studies, Fleming reported widespread professional and public endorsement of the central idea of flexible course selection. He also pointed out the stability of course and program selection among students, who had freedom to make dramatic changes if they so wished. Fleming and his colleagues also discovered considerable stability in teaching methods, course content, and within-school organization. In one of the HSl studies Laxer was alerted to the apparent inflation of marks, which even at that early stage was beginning to signal what people have called the decline in standards, which in fact may be nothing more than a decline in the value of marks while the standards of performance remain stable.

In spite of the progressive climate of the late 1960s, some officials intended that a standardized test program for the province of Ontario would replace the Grade Thirteen examinations. The Ontario Mathematics Achievement Test and the Ontario Physics Achievement Test were developed at the time for that purpose. The fact that both these tests have been chosen as

the best available instruments in their fields to measure student achievement in Ontario in 1976 is significant testimony to the stability of those two courses during the period in question.

While the critics have focussed on what they believe to be deterioration of basic education and student performance in the fundamentals, a very real change and perhaps a much more significant change has taken place almost without notice. Fleming reports a rapid increase in the number of students in secondary schools. If we examine figures at five-year intervals from 1945-46 to 1970-71, we find the number of students to be 120,000, 131,000, 175,000, 263,000, 419,000 and 557,000 respectively.¹ In 1971-72 it moves up again to 575,000. This represents a five-fold increase in gross numbers of students, and a change in the proportion of secondary students from about 40% to about 80% of potential students. It is true then that about half of students in our secondary schools today would not have been students even fifteen years ago. The schools have been required to cope with increased numbers of the type of student they have traditionally served, and have also had to provide service (and perhaps a new kind of service) to as many people again who in the past would not have expected secondary education. There have been profound ramifications to this significant change in education, the first and most obvious of which has been the cost to the community; a second has been the change in what can reasonably be expected of the average student.

If the new students in the secondary school are expected to perform as well as the traditional secondary school student, there is a problem. It is likely true, however, that the general public expects these new high school students to be of lesser academic ability, and their expectations therefore may be more modest and more realistic. In the United States, for example, there appears to be no particular blame assigned to the schools or programs for the apparent small declines that have shown up across a broad array of subjects (Gallup). The Project I report (Lo Presti) covers related issues as they apply to Ontario citizens, and any differences which emerge will require careful

consideration. Another U.S. report, Achievement Test Score Decline: Do We Need To Worry? (Harnischfagen and Wiley) leaves open the possibility that the apparent declines may be attributable to changes in the composition of the group of test takers. This point is perhaps even more applicable in the Ontario setting where the expansion in the numbers and proportions of students in the upper grades has been more recent than in the U.S.

The dramatic changes in numbers and proportions evidenced in the secondary school are even more pronounced at the postsecondary level. Again according to Fleming, the number of postsecondary students attending Colleges of Applied Arts and Technology expanded from 11,000 in 1967-68 to 37,000 in 1971-72. The university population expanded from 19,000 in 1952-53 to 117,000 in 1970-71. This seems roughly an eight-fold increase in postsecondary students from 1952-53 to the present, and the accommodation of postsecondary instructors to the changing composition of the student body appears to be as yet incomplete. Expectations on the part of postsecondary instructors are very high indeed by comparison with the performance levels the students evidence and with the expectations of secondary school instructors. Such expectations may in part be explained by the changing composition of the student body, and although this issue and related issues are important to the Interproject Analysis, they are of central concern to Project III (King et al).

Whether or not there are changes in the performance of students in Ontario Grade Thirteen is covered in the Project II report (Traub). Regardless of the main causes of changes, if any, which have arisen, there remains the central problem of interproject analysis--namely, which topics evidence gaps and which display duplications.

A number of studies beyond Ontario's borders have provided data that speak to some issues of interest in this study. One study focusses on articulation between colleges of liberal arts and secondary schools. The report, entitled A New System of

Education (Blanchard), examines the extent to which repetition occurs in selected courses, the reasons for this repetition and its effect on the performance of college students. Using a strategy similar to, but not nearly so comprehensive as, that employed in the Project III study, Blanchard was able to detect clearly defined areas of duplication in Mathematics, Social Science, English and Science. It is significant that he did not identify gaps, but gaps are less likely to be perceived when student performance is neglected. In the Eight Year Study (Aitken), there is a detailed account of a large-scale American attempt to test the hypothesis that the rigid set of entrance prerequisites for university prevailing in the United States at the beginning of the 1930s could be by-passed in favour of judgements made of the students by their own high school teachers. The extent to which the positive findings of this study (that the hypothesis is accepted) apply in Ontario today can be determined only on the basis of cooperative judgements of a variety of experts that take into consideration the main characteristics of the sample schools and universities of the Eight Year Study and the corresponding institutions in Ontario today. The judgement of the present investigators is that for most secondary schools and universities in Ontario, and for most purposes, the secondary school staffs are capable of the task of screening students for postsecondary education.

One other set of studies, the International Study of Educational Achievement (Hussein), provides data on the standards of performance of superior students who attend composite schools. This is apparently a current concern in Ontario because of the rapid increase in the proportion of students attending senior secondary schools and postsecondary institutions. If there is a tendency on the part of teachers to teach to the "lowest common denominator" of students, then it may be true that the better students are deprived of the kind of excellence in teaching that would have been provided for them had the expansion of the student population not taken place. Admittedly the IEA studies do not deal directly with the situation in Ontario, but they do compare the top 4% of students in the final grades of high school

in countries where there are composite schools with the top 4% of students in the countries where the academic elite receive secondary and postsecondary education that is specifically designed for them. The conclusion of Hussein on this point is that there is no apparent loss to the elite group. This conclusion foreshadows some of the findings of the Project II study (Traub).

There is one other study which, although it deals exclusively with mathematics, reveals both a professional and an academic point of view that in some respects coincide with the views of the general public. The study, Mathematical Sciences in Canada (Coleman), sampled opinions of mathematics professors and mathematics "users" about their own mathematical preparation in secondary school and university. The point of interest for the present study is the widespread dissatisfaction with the teaching of mathematics felt by this especially knowledgeable group of citizens.² The nature of the dissatisfaction is discussed: in some ways it is in conflict with the hard data provided in the chapters which follow (see especially Chapters Four and Five) and in the Project II report.

Another group of knowledgeable observers express sentiments similar to those of the mathematicians reported above. An OSSIF study entitled A Revised Interim Report on the Role of the Secondary School Study (Head), based on data collected from 34,000 Ontario secondary school teachers, states: "Since so many people feel that the communication and computational skills are lacking in our students at all levels and years of schooling, an emergency interim measure must be adopted. We must add extra classes on to English and Mathematics programs."

2. RELATIONSHIPS OF PROGRAMS TO TESTS

The program of studies for Ontario students in their last year of high school is flexible in that each student is free to choose courses according to his/her own requirements. Furthermore the Ontario programs are flexible in content according to the requirements of the school community or even the individual class. This latter variety of flexibility may be much more limited in practice than Ministry guidelines permit but it is not likely to be so severely limited as it was before 1967, during the years of the Grade Thirteen departmental examinations. After high school the Ontario student may begin postsecondary education, and is free to choose from a wider variety of courses and programs than were available in secondary school. Furthermore, within particular subject domains there is a wider variety of content offered than was the case in the final year of high school. Under circumstances of such apparent variation the question arises as to whether any single test or battery of subject tests can be expected to give a fair estimate of each student's competence on leaving secondary school.

During the period when Grade Thirteen departmental examinations were administered across the province, there was some general confidence that the examination marks reflected student competence and the likelihood of success in postsecondary education. However, the examinations depended upon province-wide uniformity of courses for their validity, and this uniformity in turn was fostered because of the great importance of the marks to the schools, teachers, and students. Both the general public and the academic community were critical of the severe limitations that the examinations placed on school programs, not only in Grade Thirteen but also in the lower grades and the non-university preparatory course. It seemed that there was general agreement that for each new generation of students there were many things to learn, far too many for all students to be required to learn them all. The thesis that schools could choose well what was important for their own students, and at the same time provide good preparation for even the most rigorous

university entrance requirements, was generally accepted; this can be documented by at least one research study, the Eight Year Study. Now Ontario schools have diverse programs (or at least are free to have them), and this aggravates the problem of finding test instruments able to both measure student progress in a given course and predict success in postsecondary courses.

In the United States the College Entrance Examination Board and the American College Testing Program have been preparing tests for such purposes. One index of their success is their widespread use. The Canadian counterpart to the American testing programs was SACU (see Traub, et al.). Although in theory it offered a solution to the testing problem arising from the abandonment of Grade Thirteen examinations in Ontario, it did not enjoy the kind of success that had been anticipated based on the American experience. Whether or not Ontario's general avoidance for a period of eight years of province-wide tests, either departmental or standardized, had any pronounced impact on Ontario students is difficult to determine on the basis of data available before the present study. The HSI studies of 1973 have shown that the kind of widespread diversity that was anticipated in the period following the abolition of departmental examinations did not materialize. The question of program diversity is covered in the Project III study of the nature of programs.

3. TOPIC SCORES VS AGGREGATED SCORES

One of the features common to standardized tests and Grade Thirteen examinations is that over the years virtually all record-keeping and reporting systems have used individual student data that are aggregated over an entire subject discipline--i.e., one mark for each subject taken by an individual student. Thus results on test items or examination questions which rely on vastly different skills, such as rote memory and complex reasoning, become indistinguishable. There is some justification for such aggregating in the case of standardized tests as opposed

to the old Grade Thirteen examinations, but the force of the argument in favour of aggregating shows a serious weakness of standardized tests.

The manner in which standardized tests are constructed tends to "homogenize" the items. This is achieved by an item selection procedure that screens out items which affect students differently from the way in which the majority of items under consideration affect them. This procedure selects items with a "family resemblance", and therefore aggregating student scores on a large number of similar items makes sense. What may have happened in the standardization process, however, is the screening out of items testing some important concepts or topics which lack the "family resemblance" quality.

In jurisdictions where either a standardized test or a Grade Thirteen type of examination has earned high status as a predictor of college success or whatever, there has been the fear that the test would determine the program rather than vice versa. In the case of the Grade Thirteen type of examination, there may be some breadth of program retained, but it is rigid across the entire province. In the case of programs determined by standardized tests there is not so much hope for breadth, and there is still the rigidity of program to contend with.

On the basis of these considerations the interproject analysis plan has involved topic-by-topic and item-by-item accounting of student performance data, wherever this is possible. Such an accounting is designed to permit the careful analysis of performance on all topics that are tested and to identify all topics that are not tested.

4. STUDY PLAN

The strategy in Interproject Analysis is to merge the data from Projects II and III in a manner that yields insights into fundamental problems of the secondary-postsecondary interface. The need for topic-by-topic and item-by-item merging has been outlined above, and what remain to be discussed are the nature of the data and the nature of the samples.

The data from Project II in most instances result from the administration of standardized tests, each of which is described in the appropriate chapter below and in the Project II report (Traub et al., 1976).

The test items that form the basis for description of student achievement possess some characteristics which require special consideration in view of the special purposes of the study. Although the item pool from which the standardized test items have been chosen may cover a broad domain of content, in fact may cover all the content of interest in the particular subject area, the resulting test may cover only a restricted range of material. The item selection procedures which are appropriate for the purpose of increasing test reliability can have the effect of restricting the domain of content. For instance, very difficult items are automatically screened out on the basis of their limited capability to discriminate among students. It is generally true that items which discriminate best have a difficulty level in the middle range; thus approximately half of the students trying such an item pass it and half fail it. The fifty per cent passing criterion is employed in test development even though all of the students taking the test item have been taught the topic and have presumably completed their learning of it. This basic fact about standardized test items is especially important for interpretation of the data in the Interface Study and it will create problems of interpretation that vary from subject to subject and situation to situation. Some tests, such as OPAT, have been designed to discriminate among students who are especially well prepared in physics and are

competing for places in university physics programs. The fifty percent passing criterion is applied to the primary group of students, and the result is a considerably lower mean performance for the less select, total group of physics students in Grade Thirteen. In other tests the focus of attention is the lower-performing student and in such circumstances the mean performance level for all students is well above 50% (i.e., French).

Another important characteristic of test items on standardized tests results from an item analysis procedure which screens out items that fail to favour the students who have tended to be favoured by the majority of the test items in the item pool. It is this item analysis procedure which gives the test items on standardized tests a "family resemblance" and ensures that test items without the family resemblance are removed.

Although the tests selected for use in the Interface Study are, by agreement of broadly representative committees, the best available from the point of view of predictive validity and content validity, and although the technical features of the tests meet generally accepted standards, there are reasons for concern on validity issues. In terms of predictive validity, the tests are better or worse depending on how well they predict future success of the students taking the tests as the students proceed to postsecondary courses. The central problem is that there are many postsecondary education courses within any subject discipline and within any particular university. Beyond that there are many universities and colleges within which students may pursue their postsecondary education. Thus the problem of predictive validity is to find the right criterion of success from among the multitude available. Beyond the problem of between-course variation and between-institution variation there is the other problem of year-to-year variation in the criterion within a single institution and within a single course. Thus there are many criteria available to use as a basis for assessing predictive validity.

The issue of content validity is complex also. Content in a subject area in the last year of high school can vary from course to course within a high school and it can vary from high school to high school. The chapters which follow discuss content variability, as does the Project III report, and they give support to the claim that there is some variability. In spite of these problems the tests used in the Interface Study are the best available, and are good tests by the standards that prevail at the present time in North America and the western world. In spite of the high quality of the tests it is easy to establish that they are of limited use for some of the particular purposes that are to be served by the Interface Study. What appears to be needed to describe student capability in the wide variety of topics covered in the multitude of courses at the interface is a test, or a test item, or an indicator, for each of the topics of interest. Such tests have a label--criterion-referenced tests--and they have a short history of theoretical development. In practice, however, there are none available for this study. The procedure employed in the Interface Studies has been to use the best available standardized tests as both standardized tests and CR tests. They are standardized when aggregate scores are used and they simulate CR tests when test items are used individually or in clusters according to some topic name. Thus by maintaining test item data on an item-by-item basis rather than aggregating to some total score it is possible to study each topic separately. The data then can be related to the Project III topic data and after some analyses the gaps and duplications of concern can be identified.

Besides the item performance data, Project II provides test appraisal data which result from the completion of Test Appraisal Inventories for each test by designated teachers who either teach a course under consideration or receive students from such courses. The plan of the inventories is elaborated in the Project II Report (see Traub et al.), and for the present purpose a brief account may suffice.

For each test in Project II there is a companion inventory. Most of these require the respondent to react to each item on the test (the exceptions are the inventories for both first language tests and for the test of English as a second language). In the case of secondary school teachers the category options include:

A. Old knowledge that students should have on entry to the course.

A1. This knowledge is not reviewed in the course.

A2. This knowledge is reviewed in the course.

B. New knowledge that all students are expected to learn in the course.

C. New knowledge that some students are expected to learn in the course.

C1. Only 1% to 25% of students should learn this.

C2. Only 26% to 50% of students should learn this.

C3. Only 51% to 75% of students should learn this.

C4. More than 75% but not all students should learn this.

D. New knowledge that no student is expected to learn.

This detailed classification complicates analysis by comparison with simpler forms of "opportunity to learn" indices which have been employed in other studies. However it was in fact the requirements of interproject analysis which dictated the more extensive breakdown.

If a topic is assumed to be known on entry to the last secondary year and not reviewed at that level, there is the possibility of a "super gap" developing, so that students perform badly at secondary exit in skills which they were taught in earlier years. Also these fine distinctions are required to show duplications--when the post-secondary instructors take university or college time to review material that is very well known. The breakdown of the C category seemed to be a good idea in view of the program flexibility that is permitted in Ontario.

The postsecondary form of the Test Appraisal Inventories is virtually the same as that of the secondary form with the exception that the C and D categories are replaced by "C. Other". The distinctions among the subdivisions of the C category, and between C and D, would be of significance only if the interface between freshman and sophomore years were also being studied.

The Project III data of significance in interproject analysis are the summaries of instructors' responses to the parts of the Course Description Questionnaire which can be directly related to the test data of Project II. In most instances this is the part of the questionnaire which requires the instructors to judge the performance level of their students on entry to their courses and on exit. The instructor was asked to indicate a preferred level of performance as well as the actual level. This means that for each element of the Project III Content Matrix there are eight sets of data with the corresponding provincial means and standard deviations.

The judgement of the instructors was reported with reference to different scales for different subjects, although in each case the zero end defines no competence whatsoever, and the high end defines complete mastery. Thus variation among topics can be shown, as well as differences between instructors from the two sides of the interface. The Project III report deals with these data in detail and with their relationship to other data from Project III. The interproject analysis introduces these data as

evidence of impressions or informed judgements of student competence; in cases where the topics from Project III and the items from Project II coincide, it is possible to investigate the relationships between program and student performance in some detail.

5. POPULATIONS AND SAMPLES

The nature of the populations and samples required for Projects II and III and for Interproject Analysis of necessity imposed three sets of restrictions on the sampling procedures. Project II requirements for the sampling of students are considered in the Project II report (Traub et al., 1976); they call for a group of students representative of the Ontario student population, large enough to bring error estimates to a level sufficiently low to permit decisions to be taken. Project III required a sample of programs for each subject discipline at both secondary and postsecondary levels; again the main issues are presented in the Project III report (King et al., 1976). Representativeness of the population of programs is the key issue in this case, but there are no clearly accepted statistical procedures for ensuring representativeness within clearly defined limits of error. The procedure for obtaining an acceptable degree of representativeness is described and explained in the Project III report, and no further details are presented here.

The Interproject Analysis plans for a topic-by-topic and item-by-item analysis imposed the main condition on the sampling for Projects II and III, namely that the students sampled in Project II correspond directly to the programs sampled in Project III. There would not be much hope for a detailed analysis by topics if independent samples were used to create representative provincial data, and then the provincial data from students and programs were merged. Such a scheme would work if there were in fact a provincial program that was rigidly followed throughout the province. But that was assumed not to be the case, and hence

a sampling procedure was developed jointly between the investigators of the two projects such that the sample students were taught by teachers of the sample programs.

The students of interest to the study consisted of all those who were eligible for a Secondary School Graduation Diploma or a Secondary School Honour Graduation Diploma in June 1976, assuming successful completion, by that time, of all courses in which they were enrolled. The programs of interest included all those in English, Mathematics, Physics, French, Anglais and Français taken by the sampled students.

Some important exclusions from the populations were made in the interests of efficiency. They were the students in private schools, special vocational schools, hospital schools, schools for the deaf and blind, YMCA schools and alternative schools. Also excluded were the group of students who are Anglophone and in attendance at bilingual schools, and their programs. The effect of this decision was to remove some of the students from the Anglophone population, but the proportion was so small--of the order of 1.5 per cent--that the decision was readily endorsed. The removal of the Anglophone students in bilingual schools made possible the inclusion of Francophone students who attended secondary schools in which instruction was carried out either exclusively in French or in both French and English, and who were in classes receiving instruction in French. Without the exclusion of these Anglophone students there would have been schools in which testing had to be carried out in both French and English, and interviewing and questionnaire administration done in both English and French; the amount of disruption that would have been entailed could conceivably distort the Francophone data.

The primary sampling unit for the Project II and Project III samples was the school, even though Project II ultimately deals with students and Project III with programs. By drawing a probability sample using schools as a unit it was possible to select samples representative of both students and programs. Needless to note, the representativeness must be interpreted as

within statistical limits. The main effect of the randomization was to remove bugs in the sampling process.

Primary sampling units for the Anglophone population were defined on the basis of geographic region, size of school, and the ratio of the number of SSHGD to the number of SSGD students in the school. The seven regional strata were as follows:

1. Metropolitan Toronto (excluding Mississauga)
2. Hamilton (Hamilton was treated as a separate region because it is unique within the province in having separate schools for Grade Thirteen students.)
3. Northern Ontario
4. Urban Western Ontario
5. Urban Eastern Ontario
6. Rural Western Ontario
7. Rural Eastern Ontario

Details on the number of strata and the number of cells in the sampling design are provided in the Project II report. For the present purpose it is sufficient to state that the sample is clearly unbiased because of random selection, and that it is representative within statistical limits.

In the Francophone sample only one stratification variable was used--namely SSHGD to SSGD ratio--and a much higher proportion of the Franco-Ontarian schools was drawn. If the proportion of the Franco-Ontarian population sampled had been the same as that for the Anglophone sample, there would have been far too few schools involved to provide any confidence in the representativeness of the sample. The sampling plan chosen

involved stratification of the Francophone schools into 15 cells with one school chosen for each cell.

The above considerations led to a sample size of 75, with 60 Anglophone schools and 15 Francophone schools. Details about the within-school sampling of students and the within-institution selection of teachers and their target courses is of greater interest for the Project II and the Project III reports, which in each case provide full documentation.

6. METHOD OF ANALYSIS

The basic plan for data analysis was the topic-by-topic and in some instances item-by-item study of Topic Tables (an example of such a table is presented as Appendix 4C to Chapter Four of this report). Each topic table summarized both Project II and Project III data for a particular topic. In cases where each test item corresponds to one and only one topic there is a good opportunity to capitalize on the merging of three types of data in one table. The student capability ratings from Project III for both secondary and post-secondary levels can be compared, as can "actual" and "preferred" ratings. The Project III data can also be compared to both the "opportunity to learn" data from the Test Appraisal Inventories, and to the observed student performance--i.e., the percentage of students in the sample who correctly answered the test item in question. Each discrepancy and each combination of discrepancies can be of interest in interproject analyses as they lead to the identification of gaps, duplications or anomalies.

Up to this point, we have not considered multiple mappings--i.e., topics with more than one item, and items with more than one topic. Neither has there been consideration of options for analysis where the available data conform to a different pattern than that presented. Each subject discipline presents unique problems for analysis and unique opportunities to

describe the actual situation at the interface. Discussion of these problems and opportunities is pursued in the subject chapters.

7. GAPS AND DUPLICATIONS

In its simplest form a gap arises when students have not yet learned some material, topic or item which is assumed by postsecondary teachers to be known by the students on entry to their courses. A duplication arises when students have learned very well some material which these same teachers assume to be unknown or inadequately known on postsecondary entry. One way to describe gaps and duplications of a more complex variety is outlined in Table 2.1, which dichotomizes each of three variables and then classifies the topic or related test item according to the resulting eight categories.

The simultaneous gap and duplication which is shown as condition 7 may be considered an anomaly, but regardless of what it is labelled there is need for further consideration between secondary and postsecondary officials responsible for programming. Particular disciplines seem to require different sorts of attention to the special nature of their gaps and duplications. Although the Table 2.1 classifications are most suited to the analysis of Grade Thirteen Mathematics and Physics, there is an added problem with the form of the data. In the case of data pertaining to topics, teachers were not asked to specify whether or not a topic was taught, but rather how well the students were perceived to have learned the topic. This amount of "learning" was recorded on scales which left the cutting points as to "adequacy" of learning to be determined arbitrarily from subject to subject.

The item-by-item analysis in both Grade Thirteen Mathematics and Physics follows the Table 2.1 categories reasonably well, except that in these cases the "whether or not

taught" is subdivided by the ABCD classifications outlined earlier. Again arbitrary cutting points have to be established that divide the items into "insufficient secondary coverage" and "sufficient secondary coverage" in the case of gaps, and "too much postsecondary coverage" and "sufficient postsecondary coverage" in the case of duplications.

There is one further type of gap that requires special consideration. It arises when a topic is not taught or reviewed in Grade Thirteen because it is assumed to be already known. Many such topics are not covered at postsecondary levels either, in spite of the failure of a high percentage of students to show mastery. Such gaps could be categorized simply as 1 in Table 2.1, but are different in nature from gaps of this type which originate at the Grade Thirteen level, and need separate discussion.

The complexities of gaps and duplications are not yet fully explained, because in the language domain there is not the close correspondence between topics (Project III) and items (Project II) that there is in the case of mathematics and physics. For some topics there are as many as twenty and thirty test items while for others there is none. Furthermore the items which do pertain to topics are in many instances related to a number of topics. Because of this increased complexity in the language area it has been necessary to consider special kinds of gaps, duplications and anomalies for each subject area.

Although the Topic Table was the basic unit for assembling or merging data that had been classified by topic or item, the data are presented in the subject chapters in more summary form. In some chapters it has been possible to provide one summary table which provides individual topic or item data on a single line; hence data on the entire program can be presented in the one table with possible gaps, duplications and anomalies identified for further analysis.

In cases where further analysis was indicated, we have provided more detailed information from the Topic Tables, and in one case engaged in the analysis of some case study data. These latter data are derived by replication of the general tabulation procedures for some Project II data with only five target institutions--three secondary schools and two universities. The case study institutions were selected so that one university was coupled with two secondary schools in the immediate vicinity of the university, and a second university was coupled with the one sample secondary school in its immediate vicinity. There is no suggestion that these are "typical" universities or typical communities. They happened to be available in our sample and were the only ones available that held any promise of integration across the interface through geographic proximity.

The rationale for constructing case study Topic Tables involved a number of factors. In planning the study there was the possibility that local gaps and duplications might arise where provincial data did not detect them, and conversely that local gaps and duplications might not exist although they were evidenced in the province-wide picture.

In the Grade Thirteen Mathematics chapter there are some case study data presented; they are of minimal interest. In the remaining chapters the case study data conformed to the provincial picture so closely that there was no point in reporting them.

FOOTNOTES

¹ These data are taken from page 35 of Fleming, Education: Ontario's Preoccupation.

² The point is revealed in tabulated data on page 50 of Coleman et al., op. cit.

REFERENCES

- Aikin, Wilford M. The story of the eight year study. Harper and Brothers, 1942.
- Blanchard, Everard B. A new system of education. ETC Publications, 1975.
- Coleman, A.J., Edwards, G.D., and Belcher, K.P. Mathematical sciences in Canada. 38th Science Council of Canada background study. Ottawa: Science Council of Canada, 1976.
- Fleming, W.G. Education: Ontario's preoccupation. Toronto: University of Toronto Press, 1972.
- Fleming, W.G. The individualized system: Findings from five studies. Toronto: The Ontario Institute for Studies in Education, 1974.
- Gallup, George H. Eighth annual Gallup Poll of the public's attitude toward the public schools. Phi Delta Kappan, October 1976, 58, 2, p. 191.
- Harnischfeger, Annegret, and Wiley, David E. Achievement test score decline: Do we need to worry?. Chicago: CEMREL Inc., December 1975.
- Head, Jim. A revised interim report on the role of the secondary school study, Summer 1975, p. 12.
- King, Alan J.C., et al. Continuity and diversity of courses. The secondary-postsecondary interface project III: Nature of programs. Kingston: Queen's University, 1976.
- Laxer, G., Traub, R.E., and Wayne, K. The individualized system: Student social and achievement patterns. HSI Studies. Toronto: The Ontario Institute for Studies in Education, 1974.
- Lo Presti, Michael, A., et al. Secondary-postsecondary interface study, project I. Roles and responsibilities of the secondary and postsecondary institutions 1976. Toronto: Canadian Facts Co. Limited, 1976.
- Traub, R., Wolfe, R., Wolfe, C., Evans, P., Russell, H.H. Secondary-postsecondary interface project II: Nature of students. Toronto: The Ontario Institute for Studies in Education, 1976.

TABLE 2.1

BASIC TYPES OF GAPS AND DUPLICATIONS

Condition	Taught at sec. level	Taught at postsec. level (Mastery not assumed)	Student Mastery to the level desired by postsec. instructors	Student Experience	Program	Student Mastery
1	No	No	No	gap	gap	gap
2	No	No	Yes		gap	
3	No	Yes	No			gap
4	No	Yes	Yes	dup		
5	Yes	No	No	gap		gap
6	Yes	No	Yes			
7	Yes	Yes	No		dup	gap
8	Yes	Yes	Yes	dup	dup	

CHAPTER THREE

ENGLISH*

1. INTRODUCTION

Although it is not the purpose of the interproject analysis of English to explore and report upon the history and the philosophy of the teaching of English or the present situation in the schools (matters fully reported in the Project III analysis), a brief excursion into that situation is necessary to explore and explain the specific limits which must be placed upon comparison of data between Projects II and III and any inferences that may be drawn from that comparison.

Specifications for Project III ("Nature of Programs") call for a careful examination of present courses at the interface: Grades Twelve and Thirteen, and the first year of CAATs and universities. Project II ("Nature of Students") examines the performance of students at the interface whether they happen currently to be taking English or not. Hence a sample of Grade Twelve and Grade Thirteen students has been drawn from all students eligible for the SSGD or SSHGD without regard to whether they are taking courses such as those described in Project III.

*(This chapter was prepared with the assistance of Sheila Hambleton, an administrative assistant in the Curriculum Department of the Toronto Board of Education.)

The problem of English course sequence at the interface is further complicated by the very extensive range of course offerings from school to school in the senior division. There is a considerable variety in approach, emphasis, and content; while some courses continue to stress the traditional "40:60" balance between composition and literature, many others do not. One purpose of the Project III study is, through examination of secondary school calendars and responses to the Course Description Questionnaire, to document some of this variety. Details and observations are contained in the Project III report.

In that report, data concerning course aims, objectives, and teacher expectations are based on course sequence: General and Advanced Level Twelve to CAAT, and Grade Thirteen to university. In order to determine relative emphasis and expectations for student performance within that sequence, the Project III team, with the advice of teachers from the three types of institutions embraced by the study, designed a Course Description Questionnaire (hereinafter CDQ) listing twenty possible course aims and forty-four objectives and sub-objectives in a herculean effort to comprehend the vastness and amorphousness of English programs: literature, language, and media studies.

The relevant specifications concerning assessment of student performance (Project II) were as follows:

Achievement test scores: First Language

For all students (completing SSGD or SSHGD) a general first language achievement test including all major components of literacy (e.g., written prose composition) at a level appropriate to those undertaking university or college education.

Revised Specifications,
"Nature of Students"

It followed that testing would not necessarily be specific to course objectives, but instead would embrace, as far as available instruments allowed, measures of general English competence.

It was expected that test objectives related to "literacy" would be found among course objectives for English, and this expectation was subsequently validated by data from both Project II and Project III. However, by no means all such objectives, particularly those concerned with literature, could be embraced by the testing program.

During the period of test selection and Course Description Questionnaire (CDQ) design, constant dialogue was maintained between Project II and Project III teams and Advisory Committees in order to achieve the best match possible between course objectives and test instruments. Three members of the Project II English Advisory Committee attended Project III Advisory Committee sessions, and the coordinator of the first acted as liaison to the second.

Because of the constraints in testing and the differing specifications of the two projects, relatively few of the objectives discerned by the Project III team and subsequently listed in the CDQ were measured by the Project II tests. Specifically, eight of 44 were tested, some by more than one measure (see Tables 3.1 and 3.2).

Project III findings are reported as the consensus of views of specific teachers about their courses and expectations for student performance, whereas Project II reports performance data for groupings of students electing to proceed:

- from Grade Twelve to employment;
- from Grade Twelve to postsecondary education;
- from Grade Twelve to Grade Thirteen;
- from Grade Thirteen to employment;
- from Grade Thirteen to postsecondary institutions.

Performance of the second and fifth of these groups is considered in the present Interproject Analysis.

2. TEST SELECTION AND COURSE OBJECTIVES

2.1 Advisory Committee Tasks

Advisory Committees to both projects were faced with very different tasks. The Project III committee had the unenviable task of determining what in fact constituted "English", and rendering this definition in terms of aims and objectives precise enough to have substantial meaning, sufficiently numerous to "cover" the range encompassed by a wide variety of English courses, and at the same time not in such numbers as to intimidate the hardy souls responding to the questionnaire. Meanwhile the Project II team had somehow to develop an adequate definition of postsecondary literacy (a task not yet undertaken by the Globe and Mail, though a frequent subject of its headlines and editorial discourses), and as well to discover or invent tests and testing procedures which would be perceived by the public, those educationally "literate" and those who are not, as providing a reasonable description of the language performance of SSGD and SSHGD students entering postsecondary institutions.

2.2 Test Characteristics and Cautions about Interpretation of Data

The Project II Technical Report on the Tests of English for Anglophones (Section 1, "Test Selection and Test Content") reviews the decisions made by the English Advisory Committee and a number of important issues of test design and test limitations.

One particular feature of test design does require brief rehearsal in the present report. Given that a sample of all students in all programs in Grades Twelve and Thirteen was tested

for "literacy" in English, it was essential that the tests selected or constructed must be designed to spread students over the range of scores on the test--i.e., tests must have a maximum discrimination value to distinguish between performance across a very wide range of student ability. In short, what was required, even had there not been other forceful reasons for that selection, was a standardized test of English rather than a criterion-referenced test designed to measure achievement of relatively discrete objectives. Thus the Advisory Committee was necessarily led to consider and to recommend subtests from the testing program of the Service for Admission to College and University (SACU). Items in the SACU bank are systematically constructed to produce a wide distribution of scores rather than a "bunching" of scores above and below some pass ("literate")/fail ("illiterate") line. All reporting of test scores and of the holistic scores on essays, then, is provided in percentiles to reflect the distribution of scores and to underline differences among the different streams of students.

The consequence is that, first, scores for the weakest students will appear very low indeed and scores for the above average student will not appear as exceptionally high when one looks at the scores alone, and the casual reader may assume that the heavens are falling or have fallen. This is not, however, a conclusion which the type of test or scoring strategy can support. Also, it should be noted that the test items drawn from the SACU pool were originally designed to test only university-bound students; consequently the distribution of scores is certain to be unflattering to students in a lower grade, particularly in the general stream.

2.3 Restriction of Test Objectives

The three constraints that required severe limitation of objectives to be tested were (1) the requirement that "literacy", inclusive of a sample of writing, was to be tested, (2) the

current state of the art of testing reading and language, and (3) time.

Time first of all did not permit the development, pretesting and refinement of new testing instruments prior to the May 1976 testing date. Consequently the Advisory Committee made its selection, with some reservations, from extant standardized instruments previously used at the Grade Thirteen-university interface. Additionally, it selected for the essay a modification of the holistic scoring method described in Britton, Martin, and Rosen (1966) in order to achieve satisfactory reliability.

Second, the amount of testing time available for English was much too short to permit a wide range of tests. Testing was restricted to a single school day, and, with all students in the sample expected to take some form of English test, it was necessary to use split tests and to restrict the essay test to a subsample of students. The same time constraint caused the testing of writing to be limited to a single composition.

2.4 The Essay Test and Selection of Writing Mode

The possibility of requesting two samples of writing was discussed at considerable length, as a very good case from the standpoints both of validity and of reliability could be made (Britton et al., 1966; Godshalk, Swineford, and Coffin, 1966) for increasing the number of writing assignments per student to be scored holistically. The limited testing time available, however, precluded the assignment of two or more compositions. Also considered was the essay plus a writing task such as a business letter, the latter to be scored on a three-point scale as was done in the first round Writing Assessment in the American National Assessment of Educational Progress program. The latter exercise was perceived as having pertinence particularly for students entering CAATs or the job market. (As noted in the Test Appraisal Inventory report from Project II, some support for the "writing task" option was received from CAAT instructors.

Unfortunately time, the perceived need to sample all students on the same basic tests, and logistical problems of moving students about for a variety of test packages each with different instructions forced the Committee to drop consideration of this alternative.)

At the same time, to increase the reliability of holistic scoring, it was agreed that writing must be restricted to a single mode (expository--argument or viewpoint) though a choice of topics was provided. There was general concurrence from the Committee, subsequently supported by the Test Appraisal Inventory (henceforth TAI) responses (see Table 3.6.1) that this mode was most suited to testing effectiveness in presenting, organizing and substantiating a position, features deemed essential in any definition of postsecondary "literacy". Teachers were strongly agreed that the difficulty level was reasonable, and that the topics selected and the range of topics were satisfactory. There was some criticism of the restriction to a single mode on the grounds that not all students could demonstrate their best writing when confined to this mode; some might perform better in narrative, etc. However, there was little consensus concerning an alternative mode.

Between 48% and 58% (each level represented) of teachers responding to the TAI identified the chosen mode as the most appropriate as a test of student literacy. Among the remainder there was no consensus concerning alternatives. No other mode (with the exception of business communications, which was noted by CAAT instructors) received a response exceeding 15% from any group.

2.5 The Essay Test: Relation to Course Objectives, Project III

The Project III questionnaire lists under objective No. 10 "Facility in writing" in a wide variety of modes, from tasks (as in précis) to creative writing (as in poetry). Objective No. 16 concerns "argument". The restriction to a single mode for the

testing of writing (Project II) reduced significantly the basis for comparing objectives and performance in interproject analysis.

Specifically, respecting modes of writing, comparison of data is restricted to the following Project III objectives:

- 10(a). Demonstrate facility in writing in terms of planning, organization, presentation and editing: (a) the essay;

and

16. Present an argument effectively.

Achievement level on the first is derived from holistic scoring of the essays and post-secondary appraisals. The second is assessed through the positive characterization of the writing with reference to the specific criterion for argument (Appendix 3A).

Through error counts for conventions (Table 3.4), a measure has been taken of "presentation and editing" for objective 10(a), and again through assessment of positive characteristics, "planning, organization" is assessed. As well achievement of other Project III objectives is in part measured through analysis of the essays:

8. Write an effective sentence (error count and positive characterization).
- 12/13. Use words with precision and subtlety (error count and positive characterization).
11. Use language in a grammatically appropriate manner (error count).

The students' performance in written work thus was employed in as many ways as possible to obtain some measure of achievement for objectives in the Project III CDQ.

2.6 The Language Tests

Given the constraints referred to earlier, it was essential to decide upon some form of standardized language achievement test. Several members of the Project II Advisory Committee had serious reservations about the validity of such tests (reservations obviously shared by the teachers who responded to the TAI), so very careful study was made of available alternatives. All types of items employed in the SACU testing programme were examined carefully with reference to content and construct validity, and special note was made of the study (Godshalk et al., 1966) of correlation between scores on various mixtures of item types and scores on essays.

The committee's first preference, because the item type called upon the student to make a full reconstruction of a sentence, was the "Construction Shift" represented in Part 1 of the recent CTEL (Canadian Test of English Language) administration to university freshmen. This item type had shown a moderately high correlation (0.645) with total essay scores according to the Godshalk study. Some concern was expressed about the complexity of instructions for students unfamiliar with that type of item. The committee's second preference was for either the "Sentence Correction" item type (correlation 0.705 in the Godshalk study) or the "Usage" item type (0.707 in Godshalk, the highest correlation of an individual subtest with total essay score). Careful examination of the usage items in CTEL revealed a number of items faulty from the standpoint of current standard English. It was therefore dismissed from consideration and the sentence correction subtest from a recent CELAT administration was recommended instead.

No single item on these multiple-choice tests in fact measures a single definable skill or specific objective. All relate in a general way to the student's competence to manipulate structure, or knowledge of the grammatical and other conventions of the language, but to identify specific items with specific objectives would, even if theoretically feasible, be an

unpromising task. The best that can be achieved is a rough classification of the kinds of abilities which appear to be the principal focus of the test as a whole or of particular groups of items (though overlap characterizes almost all). Summaries in that general form (see Table 3.7) were provided for the Project III team in order to facilitate a better fit between the CDQ and the Project II testing. However, because of the construction of the language tests as a sort of general blitz on language problem areas, it is reasonable only to report the relationship between the two language subtests and the Project III objectives as follows:

Project III Objective	Language Subtest
8. Write an effective sentence	Construction Shift (items 11-21)
11. Use language in a gram- matically appropriate manner	Sentence Correction (items 22-36)

As further measures of achievement with respect to these two objectives we are able to draw also upon essay performance data, as noted in section 2.4.

The English Advisory Committee's reservations about the multiple-choice language tests included concern that they (especially the construction shift item) might be too difficult for general level Grade Twelve students, their divorce from any meaningful context, their necessary preoccupation with error and sometimes with rather minute subtleties of language, and the fact that they simply do not measure the student's language performance in real communication situations.

These reservations were echoed strongly in TAI responses (Table 3.6.2).

Note that teachers of general level Grade Twelve students did perceive the language test and particularly the construction shift item type as too difficult, an opinion reinforced by the Grade Twelve score patterns on the tests. Most teachers found, however, that the objectives measured were generally appropriate and reasonably balanced, though most teachers at all levels were insistent that language testing be supplemented by a sample or samples of student writing.

Further details of teacher reaction to the language tests, together with comments and recommendations based upon that reaction, may be found in Appendix A1, Section 2 of the Project II report.

2.7 A Note Concerning Historical Patterns of Test Results: Language Achievement and Reading

Though the items in each of the three subtests have been used in earlier test administrations (CTEL 1975, CSAT 1973, CELAT 1970), comparison of results is unfortunately impossible, and this for two principal reasons.

In the first place, the administrations of the tests were quite different. In the Interface administration, as noted above, tests were split so that no student in fact took a subtest that any former groups had taken. In that "split" administration, even for the students who took both forms, speededness (see Table 3.3) became a very significant factor in two of the subtests. As well, at least in SACU administrations, students and teachers had more preparation time with the Student Handbook to ensure familiarity with the items, and in the SACU administrations it may be presumed that students perceived direct benefits to themselves in performing well on the tests.

Secondly, the populations taking the tests in different administrations were in no way equal. In the SACU program normally only those students in Grade Thirteen enrolling in universities took the tests; for CIEL, students already admitted to university (and a somewhat self-selected portion of those) took the tests. Consequently, the only group in the Interface study for which any comparison would be meaningful, even if conditions of administration had been the same, is that group of Grade Thirteen students planning postsecondary education, and even there CAAT-bound as well as university-bound students are included.

2.8 Reading Comprehension Test

Surprisingly, the Advisory Committee found its most difficult task to be the selection of an appropriate test of reading comprehension. The commercial products generally ran from poor to bad on the balance between inference items and items concerned with literal comprehension (i.e., ideas or details present in the passage expressed in somewhat different wording in the question). When tests from the SATO and SACU programs were examined carefully, it was found that inference and "main ideas" items had better representation in the SACU tests. However, the reading passages were generally very dense, often rather dull, and some were rather technical in content. Efforts were made to find one complete subtest from a past SACU administration that contained four passages varied in style and sufficiently high in interest that even weaker students would at least take a fair run at them. The Committee's greatest concern was that if Grade Twelve students found the tests overly difficult many would simply give up. Consequently any apparent measure of their reading competence would be invalid.

A number of objectives in the Project III CDQ relate directly or indirectly to reading ability. A good number of these relate fairly exclusively to the study of literature, an area not tested in Project II. Others (e.g., No. 2, "Understand the subtler nuances of language"; No. 3, "Distinguish between

essential and inessential information"; and No. 7, "Apply flexibility in the speed of reading appropriate to content and purpose") doubtless are related in some manner to the student's reading performance in any reading test; however, no test items deal with any of test scores skills directly, so no attempt is made to relate test scores to these CDQ objectives.

Reports on decisions concerning reading tests, and on the limits of what those under consideration would be perceived as testing, were sent to the Project III team. The final memorandum appears following Table 3.7 in this report.

The specific Project III objectives, then, which it can be claimed the reading test measures are these:

1. Comprehend a variety of materials (essential meaning and significant details).
5. Apply inferential skills.

3. PROGRAM EMPHASIS

One measure of the appropriateness of the tests selected is their conformity with objectives considered important by all or almost all of the teachers of English at the senior secondary level and CAAT and university instructors of first year students.

Pages 8 and 9 of the Project III CDQ provide a section concerning weight (amount of emphasis) given some twenty general aims, and pages 10 and 11, concerning course organization, ask for the proportion of instructional time given to particular aspects of program. Similarly, along with questions about expectations concerning performance on objectives tested, the Project II TAI asks what emphasis is given these in the program.

We shall deal with the expectations issue (some TAI questions and chiefly the "Objectives Response" section, pp. 13-16 of the CDQ) in the context of test data in the following section; presently we wish to examine teacher perception, as expressed in weight given in courses, of the importance of those objectives referred to in both the CDQ and the testing of language performance in Project II.

Because no performance data are available, many aims and objectives, as noted earlier, cannot be included in the Interproject Analysis; specifically among these exclusions are all course components and objectives dealing principally with literature, with an oral language program, drama, and media studies. The reader may observe with some dryness that not very much is left. That may be true, but what is left is considered quite important by most parties.

Therefore, from Section 3, "Aims of the Course", in the CDQ the following were selected from the twenty statements of aims provided:

Improve student's reading ability (No. 13)

Promote fluent and grammatically acceptable written English (No. 11)

Develop student's ability to organize and integrate ideas and materials (No. 8)

These aims appeared to fit best with achievement measured in Project II by the reading and language achievement tests and the essay.

For each of the twenty course aims, teachers were asked to respond on a scale of 0 - 3:

0 - No emphasis

1 - Very little emphasis

2 - Moderate emphasis

3 - A great deal of emphasis

Response for the three aims presently under consideration is plotted on a graph using mean scores of responding teacher groups on a scale of 0 - 3 (see Table 3.8.1).

Promotion of fluent, grammatical English and of the ability to organize and integrate ideas and materials obviously ranks very high in the priority of aims among all teachers of English, secondary and postsecondary. Mean scores of over 2.5 on a 0 - 3 whole number scale indicate that at least half the teachers responding rated the aim at "3"; with mean scores of 2.8 or 2.9, almost every teacher rated the aim a 3. One may say from this perception of emphasis that there is very little difference among teachers in the three institutions concerning the importance of these two aims.

Improvement of the student's reading ability receives unequal weight. For university teachers it is noted almost as high (2.7) as the others, but its weight falls to 2.3 or less for all other teacher groups, its lowest weighting coming from Grade Thirteen teachers of English (1.8).

This anomalous pattern of weight, notably between Grade Thirteen and university teachers of English, deserves further attention with reference both to information on course descriptions (Project III report) and to other indices from the CDQ, TAI and student performance data. One might well expect the close and high ranking of the other two aims with some for reading, but the high ranking of "Improve student's reading ability" in university requires some explanation.

One possibility is this:

Teachers in secondary schools and teachers at CAATs, especially those familiar with remedial programmes, may be aware of, whether or not they use, packaged reading programmes designed allegedly to "improve" reading ability in some direct way. One may recall SRA labs, the "Controlled Reader" and quite a variety of other sequential packages. Perhaps these teachers are in effect saying, "No, though through the course we pay some (or a lot) of attention to reading in a variety of ways, we do not include a formal component aimed directly at reading 'improvement'." University respondents, on the other hand, less likely to be familiar with such packaged programmes, would perhaps consider the improvement of students' reading ability as a fairly direct consequence of their accent on literature and on resource materials related to the study of literature.

The anomalous pattern could then be largely explained by differing perceptions of what constitutes the sort of instruction likely to produce "improvement" rather than by any particularly significant difference in emphasis in course content or methodology.

The following section of the CDQ (4. Course Content and Student Competence) begins with a section entitled "General Organization of the Course". Here, in the first question, teachers were asked to divide the percentage of their formal instruction time among literature, media studies, and language skills. Further questions requested a further breakdown for literature specifically and are not our concern in Interproject Analysis.

Unfortunately for the present analysis, the version of the CDQ employed with Grade Thirteen and university teachers did not supply, for question 1, any further breakdown of "language skills", while the version for CAATs and teachers of Grade Twelve made provision under the heading "language skills" for reading, writing, and speaking. We are therefore constrained to

present the summary in three tables which relate as closely as can be managed to the aspects of performance tested in Project II. These three tables appear as 3.8.2 a, b, and c, with Grade Thirteen-university data based only on the general heading "language skills", which may or may not include in differing proportions "reading", "speaking" and "writing". For CAATs and Grade Twelves, the tables summarize responses for "reading" and "writing" separately.

When this information is compared with that presented above and in Table 3.8.1 we see striking dissimilarities between the rating given to the course aims presently under examination and the amount of time more or less directly devoted to these aims according to the responses of the same teachers.

While university teachers generally claim the highest priority for these language-related aims (2.7, 2.8, and 2.9 on a scale of 0 - 3), total course time devoted to language skills is low. 50% claim that proportion of course time falls between 10% and 20%. In contrast, 66% of Grade Thirteen teachers claim to give 21% to 40% of course time to these skills.

It is quite apparent, therefore, that (with reference to commentary on Table 3.8.1) there are certainly differences in understanding between university and Grade Thirteen teachers of what constitutes "reading instruction" and that there is some divorce for all between stated aims and actual weight, as percentage of course time, given to these aims in formal instruction.

A correlation does exist between the two tables for emphasis on reading instruction for Twelve Advanced, Twelve General and CAAT teachers, though, as noted above, the question of what does or what should constitute formal instruction here remains fuzzy.

A particularly strong concern expressed by the public and the press, and by teacher response within the Interface Project, is student writing ability at the Interface and the presumably related language skills. In that light, the very high priority given to the aims "Promote fluent and grammatically acceptable written English" and "Develop student's ability to organize and integrate ideas and materials" (which for practical purposes we are taking in the Interproject Analysis to mean chiefly in writing) in Table 3.8.1 is quite consistent. However, when we examine Tables 3.8.2 A and C, we must assume that the aim falls short in practice. 26% of university teachers of English claim to give less than 21% of course time to all language skills combined, though for Grade Thirteen this drops to 26%. On the other hand, for teachers of advanced and general level Grade Twelve students these figures for writing are 49% and 57% respectively, whereas in contrast CAAT first year English instructors claim that 51% give more than 40% of instructional time to writing.

Further information concerning the weight given reading, language skills and writing (at least of essays of the general type assigned in Project II testing) is available from the TAI. (See questions concerning course emphasis in the Technical Report on the Tests of English for Anglophones, Appendix A1 to the Project II report.)

According to that data, "heavy" to "moderately heavy" emphasis is given to specific reading, language, and writing objectives by the percentages of teachers recorded in summary in Table 3.8.3.

Response generally, as reported in Table 3.8.1 and especially in Table 3.8.3, indicates strong and welcome support and validation of the test instruments selected in Project II. Whatever reservations some teachers may have had concerning multiple-choice tests, the level of difficulty of the tests used, or the choice of writing mode, both tables indicate that the kinds of language performance tested were quite appropriate for

courses in English at the Interface. Evidently, however, the same cannot be said when we consider the responses summarized in Table 3.8.2. If we consider as an important index of the fairness of a test the time spent in teaching to that objective in the course, then we should have to say that a substantial proportion of students tested who were currently taking English courses in Twelve or Thirteen had received little recent instruction or practice related to the test objectives.

In comparing the data from the three tables, one is left to puzzle out what respondents mean by "emphasis"; it does not appear to mean percentage or proportion of time given to the objective in courses. This anomaly may be partly explained with reference to the differing perceptions of what constitutes formal reading instruction, but it is difficult to wash out differences quite so easily for language instruction and writing.

Direct comparison of Tables 3.8.1 and 3.8.3 for "Ability to organize and integrate ideas and materials" and "Writing (in the mode assigned)" reveals a very direct correspondence for all courses. "Promote fluent and grammatically acceptable written English" does not, course by course, parallel "Language" in Table 3.8.3 where, first, the objective appears to receive considerably less emphasis and, second, the pattern from course to course differs.

Reading, when the subdivisions in Table 3.8.3 are averaged, follows a strikingly different pattern. In the first table, as already noted, the difference between university emphasis on the one hand and Grades Thirteen and Twelve Advanced on the other is quite large; in the second table all three give reading approximately equal emphasis (averaged percentage is 80 for university, 82 for Grade Thirteen, 83 for Twelve Advanced) while Twelve General and CAATs give lower emphasis (77% and 76% respectively). It may be, therefore, that the apparent different emphasis on reading discussed for Table 3.8.1 is in fact not a real difference.

Which of the three tables, for reading, language and writing, presents the most accurate general picture? The chief external validator we have, from the Project III report, is information about course offerings as found in secondary school and postsecondary calendars. These data strongly suggest that Table 3.8.2, emphasis as reflected in amount of instructional time given, is the most realistic description of the situation.

We can account for some of the difference between Tables 3.8.1 and 3.8.3 on the one hand and Table 3.8.2 on the other, by differing perceptions of what constitutes formal instruction in reading. However several anomalies remain for which any explanation must be conjectural.

The following conjecture or hypothesis is proposed:

Teachers of English for the levels and courses under study are strongly influenced in responding for "emphasis" by "expectation"--expectations they have and expectations which have been strongly expressed by the public, including some very influential sectors, and the press. These "expectations" are focussed precisely upon performance in the areas of reading, language competence, and writing. Consequently, when asked "What emphasis do you give this objective?" teachers respond, "Oh, a very great deal!" and this may in turn be translated in any or all of the following ways:

"We see reading competence and/or language competence and/or writing ability as essential outcomes of secondary school English programs."

"We expect (in the sense of 'ought to find even when we don't find') these competences in our students on entry to Twelve or Thirteen or first year postsecondary, and we see these abilities as vital prerequisites for success in our courses."

"We do (at least for language skills and writing ability) take heavily into account these skills when we evaluate student performance in our subject." (Hence heavy emphasis, even where the skills are not directly taught to any large extent in the course?)

This writer finds it very hard to believe, despite a 96% response from university first year teachers of English stating that they give "heavy" or "moderately heavy" emphasis to writing in the mode assigned (exposition--argument or viewpoint), that this is in fact the case in courses in the history of English literature from Beowulf to Auden, or even courses in contemporary literature. It seems more likely that they are saying, "We expect students on entry to university to have the ability to organize and treat with some detachment or objectivity a variety of matters concerned with the essay in literature," and perceiving the skills expected for the mode assigned in the Interface testing as relevant to that performance.

Some further evidence is available in support of the third question formulated in the hypothesis. According to the recently published study The Queen's English (Norman, 1976), instructors' marks for content on final examinations in English correlated 0.79 with scores for "literacy" (pp. 27-31). The correlation of these scores with social science and physical science examination marks was markedly lower. In one sense, then, teachers of university English, and likely other teachers of English as well, can quite properly say that particular aspects of literacy receive heavy emphasis in their courses. "Heavy emphasis" in that case would mean that they receive great weight in teacher expectations and evaluation, not necessarily that instruction related to these objectives is a significant part of the formal instruction program.

If this conjecture or hypothesis is reasonably sound, then the major discrepancy between Table 3.8.2 and the others partly dissolves. As well, a gap of considerable importance is revealed: if, in these aspects of literacy perceived as important by all,

general student performance is less than satisfactory though the expectation is reasonable, probably what is required is a considerable shift in course emphasis.

How far such a recommendation should be pressed is another question or series of questions: To major revisions of course content in all levels under study? To provision of remedial programs or courses at CAATs (already a fact of life) and at universities? To revision of course content in earlier grades? Are expectations or standards really too high?

And there is the prior issue: What constitutes "satisfactory" performance and for whom? How is it to be ascertained? These questions in turn lead to a whole series of questions concerning testing and evaluation.

One focus of the Interface Study is on the question of student performance, and to this we now turn.

4. TEACHER EXPECTATION AND STUDENT PERFORMANCE: TREATMENT OF DATA

We have hypothesized that teacher response concerning emphasis on objectives is more a function of expectation than it is of course content. In the present section, for Reading, Language, and Writing, we shall examine teacher expectation in relation to test performance. As indicated by Tables 3.1 and 3.2, we have two principal sources of information concerning expectations: expectations with reference to course objectives from the concluding section of the CDQ (Project III) and response to certain questions on the TAI (Project II). Student performance data, depending on the objective, will be drawn from subtest and essay scores or from a variety of appraisals made of a sample of the essays.

For a complete report and analysis of response for all objectives in the CDQ, the reader is referred to the Project III report. For details concerning essay appraisals and a sample of the writing, the reader is referred to the Project II report, Appendix A1, Technical Report on the Tests of English for Anglophones.

As noted in the general introduction, Project III data are course-specific; thus examination of the interface lies in relationship between exit-entry expectations for (1) Twelve Advanced and Twelve General English courses--CAAT first year regular English courses, and (2) Grade Thirteen English courses--university first year English courses. Project II, on the other hand, is concerned with students; consequently test data are reported with reference to those groups planning to continue from Grade Twelve to CAATs, Grade Twelve to Thirteen, Grade Thirteen to postsecondary, and from each grade level to employment. Test performance summarized for each group appears in the Project II report; for Interproject Analysis we shall examine data for students planning (as indicated by Student Questionnaire data) to go from Grade Twelve directly to postsecondary education and from Grade Thirteen to postsecondary education.

4.1 A Note on CDQ Data: Exit-Entry Expectation Concerning Objectives

The 0 - 7 scale teachers were asked to employ is reproduced in Table 3.1, together with a summary of mean scores for those CDQ objectives examined in the Interproject Analysis. These data refer to expected competence: i.e. for entry, "your perception of the average level of competence you find students have acquired before taking your course" and, for exit, "your perception of the average level of competence you expect students to achieve by the end of your course" (CDQ, p. 12). Please note the omnipresent discrepancy between "exit" and "entry" between institutions. Leaving aside Twelve Advanced students, many of the best of whom continue to Thirteen anyway, the discrepancy

between General Twelve exit and CAAT entry falls between 1.1 and 2.0 and that between Thirteen and university--with one exception--between 2.1 and 2.7.

Because of the relative constancy of difference (Twelve General--CAAT) and (Thirteen--university), it seems reasonable to assume that each group is perceiving the 0 - 7 scale differently, rather than that there is all that much difference in perception of the students' actual exit or entry performance.

Also in the CDQ teachers were asked to give, using the same scale, preferred achievement levels at entry and exit. As might be expected, "preferred" exceeded "actual". What is of mild interest is the consistency of the difference: on five objectives (1, 5, 8, 10a, 11) examined, the differences between actual and preferred were as follows:

Exit 13	- Range of difference 1.3 - 1.7, Mean 1.76
Entry univ.	- Range of difference 1.5 - 2.0, Mean 1.76
Exit 12G	- Range of difference 1.7 - 1.8, Mean 1.76
Entry CAAT	- Range of difference 1.7 - 1.9, Mean 1.80

In Interproject Analysis we are chiefly concerned with the present state of affairs; since the amount of difference between "actual" and "preferred" at both secondary and postsecondary levels is so close to constant across objectives and levels, we may simply note this and pass on with the observation that all involved in education are agreed that, whatever the state of affairs, one always wishes it were better.

4.2 Organization of Data

Given the form in which the data are recorded and the very large amount of agreement, noted in the previous section, about the "literacy" objectives of reading, language competence and writing, it appears that the most straightforward and least redundant way of organizing data in this chapter is under the

sub-headings Reading; Language, and Writing, with treatment of each interface separately under these subheadings, rather than by discussion of all three areas of performance together first for the Grade Thirteen--university interface and then for the Grade Twelve--CAAT interface.

5. READING COMPREHENSION

5.1 The Grade Thirteen-University Interface

In section 3--PROGRAM EMPHASIS--we commented upon the appearance of a substantial difference between Grade Thirteen and university first year emphasis on improvement of students' reading ability (Table 3.8.1). However, contrary evidence appears in Tables 3.8.2 and 3.8.3, and reasons for these differences in perception have already been fully discussed.

Concerning performance expectations we observe (Table 3.1) once again what appears to be a discrepancy for the reading objectives under examination:

	13	UNIV.	
	EXIT	ENTRY	DIFF.
1. Comprehend a variety of materials (essential meaning and significant details)	4.3	2.1	2.2
2. Apply inferential skills	3.9	1.4	2.5

The previously noted fact that the difference in expectation between Grade Thirteen and university teachers on almost all objectives lies between 2.1 and 2.7 suggests that the scale was perceived differently by the two groups. In the present instance, with differences so close to the mean, it is not clear whether the difference is of any significance.

A further indicator of expectations is provided by responses to the Test Appraisal Inventory questions concerning the difficulty level of the reading test (see Table 3.6.3). On a scale of 1 (too easy) to 5 (too difficult) the mean response from university instructors was 3.29 and from Grade Thirteen teachers 3.50, both groups judging the test in the "about right" to "somewhat difficult" range. This would suggest that there is not a very large amount of real difference between levels in their appraisal of the present level of reading competence of students.

The scores for the reading subtest, with wide distribution around a mean score of 5.1, suggest that the test was at an appropriate level of difficulty for those students in Grade Thirteen planning to enter postsecondary institutions (see Table 3.9). Achievement for this group appears reasonably satisfactory, especially when one observes (Table 3.3) that the reading test appeared to be considerably speeded; i.e. students found the time insufficient to complete the test. Note the increases in "No response" for items 8 to 10 in both Forms 1 and 2. It is impossible to determine whether some students failed to complete the late items because they found them too difficult or because of time, but it is evident that the time factor did interfere to some extent with performance.

Reading comprehension in the CDQ and TAI is broken down into objectives as follows: literal understanding, main idea, and inference (TAI, Table 3.8.3.); "Comprehend a variety of materials (essential meaning and significant details)" and "Apply inferential skills" (CDQ). In order to provide a rough comparison of student performance for the CDQ objectives, the twenty items for reading comprehension on Forms 1 and 2 combined were examined to determine which of these two skills each item appeared to be testing. Twelve of the twenty items appeared to be classifiable under "Comprehend essential meaning and significant details" and the remaining eight appeared to deal either with inference or implication. As item-by-item performance data was not available for the equated results or for the different student

groups, the unadjusted score results (Table 3.3) had to be used instead as a rough index of performance for the two objectives.

The speededness of the test proved to be a serious complicating factor even in this rough measure. Unfortunately, inference items were heavily concentrated in the latter part of the test, especially in Form 1. Therefore performance on the item groups is given as the mean score for students answering the item. Given that this provides a very crude index indeed, the results are as follows for total Grade Thirteen students:

"Comprehend essential meaning and significant details"	6.7
"Apply inferential skills"	5.2

where each score is the mean raw score for those answering the item.

It appears from this information that performance on inference items is generally less effective than on more straightforward comprehension. And it is of some interest and importance to note that the difference is reflected in other data available from both projects.

According to Table 3.1, expectation across all teacher groups is considerably lower for inference than for "essential meaning", and this perception receives support from TAI questions concerning entry expectations for "literal understanding" and "inference" (see Table 3.10). Though the expression of the questions in the TAI is very general and response from Grade Thirteen and university teachers appears rather optimistic, both groups do discern a difference in ability between reading for literal understanding and drawing inferences and implications. Consistent with this perception we find (Table 3.8.3) that greater emphasis is claimed at both levels for the teaching of skills related to inference than for those related to reading for literal understanding.

This general consistency of perception combined with the performance data we have for reading comprehension in general and for reading comprehension divided between these skill areas suggests that the anomalies perceived earlier are related to differing views of what constitutes formal reading instruction rather than to actual differences in expectation or teaching emphasis. Though no one could conclude that more emphasis should not be placed on reading program generally, the evidence does not reveal that either serious gaps or duplications exist at the interface between Grade Thirteen and universities.

If a gap does exist, it likely lies in the amount of course time allocated to formal reading instruction, as indicated in the lack of fit between Tables 3.8.1 and 3.8.3 on the one hand and Table 3.8.2 on the other. What ought to constitute "formal reading instruction" remains, however, a question outside the limits of this study.

5.2 The Grade Twelve-CAAT Interface

According to Table 3.1, we find again a discrepancy between the postsecondary and secondary expectations for exit-entry, but a considerably lower expectation from both than from Thirteen-university teachers.

	12A	12G	CAAT
	EXIT	EXIT	ENTRY
1. Comprehension	4.2	3.6	1.8
5. Inferential skills	3.5	2.8	1.3

As before, we find that expectations for literal comprehension are higher than for inferential skills, and this difference is borne out by the entry expectations data in the TAI (Table 3.10). According to that information a relatively low percentage of teachers expect more than 75% of the students entering Grade

Twelve Advanced or General to be effective in drawing inference and seeing implications. At the CAAT level, fifty per cent of teachers responding find this skill developed in 75% or more of the entering students.

Both Grade Twelve and CAAT teachers perceived the reading comprehension test as difficult for these students (Table 3.6.3), with Twelve Advanced and CAAT teachers rating it about the same on the difficulty scale. Twelve General teachers saw the test as too difficult (a mean of 4.35 on a 1 - 5 scale for index of difficulty). It is evident from the test results (Table 3.9) that the Grade Twelve General teachers were the most realistic assessors of the test.

A mean score of 2.3 for Grade Twelve students proceeding to postsecondary (presumably many of the Grade Twelve General students) indicates that the test, however well it discriminated among different classifications of students, was excessively difficult for this group. The problem of speededness, already noted, likely affected performance by these students considerably. The initial fear expressed by the Advisory Committee about the density of the reading passages was well justified.

Division of the test items according to the CDQ objectives again illustrated a substantial difference in performance on items testing literal comprehension or main ideas and items testing inference. As for the Grade Thirteens, these figures are based on raw scores (Table 3.3) of students answering the items. Means are 5.4 for literal comprehension/main idea and 3.8 for inference.

In view of test performance and the opinion of teachers concerning entry performance of students in reading for literal meaning and for inference, the emphasis given these respective skills, according to Table 3.8.3, seems appropriate. Contrary to the pattern of Grade Thirteen and university teachers who stress inference over literal understanding, teachers of Grade Twelve General and teachers of English in first year CAAT programs give

the greater emphasis to reading for literal understanding and main idea.

The question raised earlier remains: What constitutes or should constitute a formal reading program? While teachers do claim reading objectives receive considerable emphasis in their programs, the data concerning formal instruction time (Table 3.8.2B) suggest that in fact not a great deal of time is given over to it, though CAAT teachers give such instruction more time than do the others. Perhaps, in view of the test results, this is because they observe that a good deal of help is required.

It is unfortunate that the reading test proved so difficult for students from Grade Twelve planning to go on to postsecondary institutions, for the test then fails as a useful indicator of student capabilities. It should not be concluded from these results that students cannot read, only that they read less well than Grade Thirteen students planning to attend postsecondary institutions. To determine the reading achievement level of students entering postsecondary institutions from Grade Twelve, a more appropriate test needs to be found or developed.

It does appear from the data that the relatively high emphasis apparently given to reading skills in first year CAAT English programs is quite appropriate and that perhaps reading should receive more emphasis than it does in Grade Twelve programs, particularly Grade Twelve General.

6. LANGUAGE ACHIEVEMENT

6.1 The Grade Thirteen-University Interface

According to Table 3.8.1, the promotion of fluent and grammatically acceptable written English is a very high priority for teachers of English in all courses at the Interface, though according to data concerning instructional time (Table 3.8.2)

many teachers give language skills little formal study. For example, according to that table, 65% of university instructors give less than 21% of their course time to language.

Most aspects of performance on the essay will be discussed in section 7; in the present section we shall examine the specific CDQ objectives concerned with sentence structure and grammar, as these seem to be the general types of language achievement measured by the Project II language tests. In commenting on these language objectives we shall refer also to error counts for sentence structure and grammar as well as other features of sentence style observed in a sample of the essays.

The following are the exit-entry expectations (Table 3.1):

	13 EXIT	UNIV. ENTRY
8. Write an effective sentence	4.6	1.9
11. Grammar	4.2	1.9

As before for reading, we observe a secondary-postsecondary discrepancy of about 2.5 on the 0 - 7 scale of competence. (See again Table 3.1 for this scale.) It was suggested that this discrepancy because of its relative constancy over objectives may be a difference in perception of the scale rather than in perception of the students, and this appears to be borne out again here when estimates of difficulty of the language achievement tests are examined (Table 3.6.2). Means on the 1 - 5 scale of difficulty are 3.21 for Grade Thirteen teachers and 2.86 for university teachers, both groups placing the tests in the "about right" range. A comparison of exit-entry expectancies for competence in skills tested in the language achievement test (Table 3.11) also supports the conclusion that both teacher groups perceive the students in much the same way. While the word "should" in the TAI question presents some ambiguity ("normal expectation" or "preferred"?) we do find a close agreement: 83%

of the Grade Thirteen teachers and 84% of the university teachers consider that more than 75% of students (exit Thirteen, entry university) should have the skills tested in the language achievement test.

Grammar

The "sentence correction" subtest, despite its title, deals chiefly with grammar items (see Table 3.7) and so is our first index of performance. We note in Table 3.9, Section (d), a mean score of 8.0 on the fifteen items, with a wide distribution. This suggests that the test was at an appropriate level of difficulty for the students and that performance was quite acceptable.

Through error counts on a sample of fifty essays we have a further index, though the sample is a very small one, of the frequency of grammar errors over 200 running words. A complete report of the error-counting procedure with appendices describing the types of errors counted appears in Project II Report (Appendix A1, Technical Report on the Tests of English for Anglophones). For the present inquiry it should be noted that the grammar error count includes both major and minor errors as defined in the once-annual Department of Education publication concerning the marking of Grade Thirteen departmental examinations. A table of counts for each of the fifty essays appears here as Table 3.4. Fractions in that table are the result of averaging the totals of two error counters.

(In considering this table and Table 3.5, please note first that the 50 essays are drawn from all the essays written by Grade Twelve and Thirteen students whether college-bound or not. Second, these fifty drawn prior to any marking do not happen to be evenly distributed across the range: i.e. what is viewed there is not proportional to the total number of essays written. These are fifty essays for which we have highly accurate scores as they have been marked by all thirty-six scorers and the scores averaged.)

The grammar error frequency for those essays scored above the 25th percentile for Grade Thirteen students going on to postsecondary education is 0.73 errors for 200 running words (essays ranked from 1 to 26 in the table). Of course it is not infrequency of grammar errors alone that renders an essay acceptable or unacceptable, and the question of what constitutes a "tolerable" frequency of error is one the reader must answer for himself; nevertheless this count may be a pleasant surprise for those who have listened with some alarm to the general hue and cry about collapsing standards.

Sentence Structure

The "construction shift" subtest, though requiring a knowledge of grammatical structures, does call upon the student to reconstruct sentences. Here the score pattern (Table 3.9, Section c) is not quite so comforting with its mean score of 4.3 on an eleven-item test.

Two, probably related, qualifications are in order. The first is that this subtest, unlike the Sentence Correction subtest described above, appears to have been highly speeded (Table 3.3); note the percentage of non-attempts for items 18 to 21 in Form 1 and 17 to 21 in Form 2. Indeed we find a very high percentage of blanks rather than wrong answers in other items in this subtest. Part of the reason may be found in the criticisms many secondary school teachers made of the "construction shift" item. They considered the instructions very confusing and were uncertain of what really was expected. A number felt the test unfair to students who, not having encountered this type of item before, had insufficient practice time with the Student Handbook. (A more detailed report on difficulties with this item type appears in the Project II appendix concerning English tests). Certainly this criticism, which some may have heard directly from students who wrote the same day that teachers completed the Test Appraisal Inventory, would go a considerable way to accounting for such a high proportion of blanks and for the appearance of speededness.

Fortunately, there are two further measures of performance with sentence structure available from the essays: the error count on the fifty and the assessment for positive characteristics of one hundred essays (inclusive of the fifty).

For the error count, sentence errors are separately classified as major or minor as described in S.4C, the former bible for marking Grade Thirteen essays. For essays above the 25th percentile (1 to 26 in Table 3.4) there is an average frequency of major errors in sentence structure of 0.37 for 200 running words. The frequency of minor errors in sentence structure above the 25th percentile is 0.46 for 200 running words. This count, certainly for major errors, constitutes a surprise indeed for those who have become convinced that secondary school graduates cannot write a sentence.

At the same time, although an essay may be free of actual errors, it may still be composed largely of rather ineffectual sentences. The original fifty essays scored by all 36 scorers were augmented by an additional fifty, ten drawn from each decile in the total score range of the essays, and these were appraised from the standpoint of the criteria outlined in Appendix A. (A complete report of this process together with sample essays appears in the Project II Technical Report on the Tests of English for Anglophones). One of the criteria was "STYLE (SENTENCE STRUCTURE)".

From the Project II Report:

Style-Sentence Structure: Serious errors in sentence structure are infrequent right across the range (see preceding "error count" report); on the other hand, attention to sentence style appears to be generally lacking in essays scored below 6.0. Though there is some variety in length, order, and complexity, it appears rather random, without consciousness of effect. Use of parallelism and antithesis, which might be expected frequently in the presentation of an argument

or viewpoint, appears very rarely except in the essays rated near the top. In these last essays there is evidence of tautness and control; in the less distinguished essays, sentences tend to ramble and insufficient use is made of means of subordinating less important ideas.

As 75% of the Grade Thirteen students planning to continue postsecondary education wrote essays which fell above 5.7, it is reasonable to say, though the number of essays sampled is rather small, that sentence structure in a fairly high proportion of these essays is reasonably good, though with considerable room for improvement.

6.2 The Grade Twelve-CAAT Interface

Expectations for exit-entry Grade Twelve-CAAT are, from Table 3.1, as follows:

	12A EXIT	12G EXIT	CAAT ENTRY
8. Write an effect- ive sentence	4.2	3.6	1.9
11. Use language in a grammatically appropriate manner	3.8	3.3	1.8

The distance on the 0 - 7 scale of competence between Twelve General and CAAT is much as we found with reading objectives, so again the question is whether this represents a difference in perception of student performance or a difference in use of the scale. Comparison (Table 3.6.2) of the perception of the level of difficulty of the language achievement tests shows something of a discrepancy (3.71 on the 1 - 5 scale for CAAT

teachers and 4.10 for Grade Twelve General teachers) suggesting perhaps that the secondary school teachers have somewhat lower expectations. Certainly their collective perception that the test was too difficult is borne out by the results. A comparison (Table 3.11) of exit (Twelve General) and entrance (CAAT) expectancies based on study of the language tests reinforces this discrepancy. 69% of CAAT teachers expect more than 75% of incoming students to have the competences represented by the test items while only 45% of teachers of Twelve General expect this proportion to have mastered such skills by the end of Grade Twelve. Conformity between Grade Twelve Advanced and CAAT teachers, on the other hand, is very close.

We may say tentatively, then, that there is a gap between CAAT and Grade Twelve General teachers' expectations for competence in fundamental language skills.

Test results, to the extent that they can be considered a valid measure of these students' potential, underline the need for considerable improvement in language competence at this interface.

Grammar

The "sentence correction" subtest (15 items) produced a mean score of 5.5 (Table 3.9 d) for students from Grade Twelve planning to enter postsecondary institutions. Performance, then, is evidently not very strong though the test was sufficiently within their reach to produce a reasonably wide distribution of scores.

The error count for grammar (major and minor errors combined) on the fifty essays produced an average frequency of 0.78 for 200 running words for those essays over the 25th percentile, though as may be seen from Table 3.4 (right hand margin) the number of essays in this particular sample that happened to fall between the 50th and 25th percentiles was only three: i.e. the chance distribution of the fifty essays selected for random scoring plays some havoc with exactitude for the Grade

Twelve-postsecondary counts. Grammar error frequency for those essays falling between the 10th and 25th percentiles is 1.2 for 200 running words. While error count figures appear to compare rather favourably with figures for Grade Thirteen students, the essay distribution suggests caution in interpreting results. Perhaps one can say, nevertheless, that the relatively low frequency of grammar errors (which include parts of speech, pronoun faults, faulty verbs, etc.) for the essays overall is cheering news.

Sentence Structure

The "construction shift" subtest was something of a disaster for the group of students planning to pursue postsecondary education from Grade Twelve. The mean score for this eleven-item subtest was 2.2 and of course the test did not provide a particularly useful distribution of scores for these students as there wasn't much room for scores to go anywhere. As with the reading test results, it should be noted that the test was simply too difficult and the conclusion that students cannot write a sentence is wholly inappropriate. In the context of the Grade Thirteen-university interface results, above, we have noted that both speededness and problems with the instructions seriously reduced the validity of this subtest as a measure of student language achievement, and we may assume the likelihood that these Grade Twelve students probably found both factors even more overwhelming.

Commentary on sentence style (see comments for the university interface, above) with reference to the essay indicated that generally in essays scored below 6 on the 1 - 10 scale, effectiveness of sentence structure was rather infrequent. As only 25% of the essays from students planning postsecondary education directly from Grade Twelve were scored above 6.1, we may say that effectiveness in sentence style, as distinct from errors, was not characteristic of most of the writing from this group.

The error count (Table 3.4) for major errors in sentence structure in essays above the 25th percentile was quite reasonable: a frequency of 0.41 for 200 running words; for minor errors, the count was 0.44. However, the caution expressed for the grammar count should be repeated: the distribution of the fifty essays gives us only three between the 25th and 50th percentile, so extrapolation of these figures is rather unsafe. The frequency of major errors jumps to 1.5 for essays between the 10th and 25th percentiles.

6.3 Summary

In general it appears that performance of Grade Thirteen students entering postsecondary education is reasonably satisfactory with respect to the basics of grammar and sentence structure. Evidently these basics are perceived as very important, though there is evidence that language skills receive little direct, formal attention in many Grade Thirteen and university programs. And it appears that there is substantial agreement between Grade Thirteen and first-year university teachers concerning actual performance levels on these skills.

For Grade Twelve students entering CAATs there is evidence of a gap in perception of performance expectation between CAAT teachers and Grade Twelve General teachers, though this gap does not extend to teachers of Grade Twelve Advanced students. Language test evidence, to the extent that the test is valid (and that may be a very small extent) for these students, suggests that there is a problem. The relatively low frequency of errors in essays above the 25th percentile, however, suggests that the overall situation on fundamentals of grammar and sentence structure is not as bad as the many critics of secondary education would claim it to be. Internal evidence from a sample of the essays indicates, however, that these students have some way to go in making sentence structure relate effectively to purpose.

Though CDQ evidence (Table 3.8.1) indicates that heavy emphasis is given to grammatically acceptable written English. Table 3.8.2 reveals that 57% of Grade Twelve General teachers and 49% of Twelve Advanced spend less than 21% of formal instruction time on writing, while for CAAT regular first year programs 78% of instructors claim to give writing more than 20% of instructional time. It would appear that more provision should be made in Grade Twelve programs (and/or sooner) for language and writing.

7. WRITING

In earlier discussion of test selection, reasons for the limitation of the writing test to a single composition in the "expository--argument or viewpoint" mode have been outlined. That the inclusion of the writing test and the choice of this mode in particular met with favour from teachers at all levels was reported with reference to Tables 3.6.1 and 3.6.2(3). All groups agreed very strongly (Table 3.6.1) that the writing assignment was at a reasonable level of difficulty for students in courses at this level, Grade Twelve or Grade Thirteen.

Tables 3.8.1 and 3.8.3 indicate that all teachers agree that writing is a very important objective in their programs. Over 90% of university, Grade Thirteen, and Advanced Twelve teachers indicate that "heavy" to "moderately heavy" emphasis is given to the mode assigned; figures for Twelve General and CAATs are 78% and 87% respectively.

However, percentage of course time actually spent on writing or on "language skills", according to Table 3.8.2, fails for most courses to reflect this emphasis. The striking exception is the first year CAAT teachers, of whom 86% claim that over 20% of course time is devoted to writing.

Besides the score pattern derived from holistic scoring of the essays, we have from Project II (for the fifty essays scored by all 36 scorers) error count figures (Table 3.4) and assessment by university and CAAT teachers of acceptability of the essays (Table 3.5). For these fifty essays augmented by another fifty drawn from the various deciles of the score range we have an appraisal of the five positive characteristics of the writing described in Appendix A. Types of error count and selection of criteria for positive appraisal were in part determined by the objectives in the CDQ; observations concerning the essays judged from the standpoint of sentence structure and grammar errors have been discussed in the preceding section in relation to the pertinent CDQ objectives.

7.1 The Grade Thirteen-University Interface

For the CDQ objective 10(a) "Demonstrate facility in writing in terms of planning, organization, presentation and editing essays (expository prose)", Table 3.1 shows the following exit-entry expectations on the 0 - 7 scale:

13 exit - 4.5	University entrance - 1.8
---------------	---------------------------

Again there is a discrepancy in the same range as for other objectives.

Examination (Table 3.12) of exit-entry expectations for the ability to write an essay in this mode reveals that 96% of both university teachers for entrance and Grade Thirteen teachers for exit claim that more than 75% of students should be able to write an acceptable essay. One may say tentatively that the two teacher groups see students' performance in much the same way but use the scale differently.

On the holistic scoring scale of 1 - 10 we find (Table 3.9) a mean score of 6.6, which suggests that in comparison with the population of all Grade Twelve and Thirteen students writing

the essay, the university-bound group performed quite well. A score of 4.6 at the tenth percentile reveals that very few of these students wrote essays that fell near the bottom in the ranking.

Holistic scoring measures the essays in relation to each other; it does not provide a definition of "acceptable/unacceptable" writing. Consequently a sample of the essays was examined from other standpoints.

Postsecondary judgements (Table 3.5)

These fifty essays marked by all scorers were sent to a sample of instructors (in English departments and outside) at four Ontario universities and four CAATs with the request that the essays be placed under one of three general classifications:

A: Acceptable quality of writing for entry to program at this institution.

R: Remedial attention appears to be required, but some qualities of the writing appear promising for success.

X: Quality of writing not acceptable, even if some remedial attention were provided, for likely success in programs at this institution.

A detailed report of the results of this classification procedure is given in the Project II Technical Report on the Tests of English for Anglophones. As already noted concerning error counts, the fifty essays drawn at random before scoring do not effectively reflect the scale. As well, it must be remembered that the fifty essays represent the work of all students writing the test whether they are exiting from Grade Twelve to work or going from Thirteen to university.

For those essays above the 25th percentile (Grade Thirteen-university), i.e., essays 1 to 26 in rank order, university response from English and other faculty members combined (a total of 312 responses for the 26 essays) was as follows:

Acceptable - 52% of responses

Requiring some remedial attention - 38% of responses

Unacceptable - 10% of responses

Close examination of Table 3.5 reveals that all essays above the 75th percentile were judged acceptable by a majority of university instructors and seven of ten between the 75th and the 50th percentile were judged acceptable. No essays above the 25th percentile were judged as unacceptable, though a substantial number below the 50th percentile were judged as work requiring some remedial treatment.

Generally it may be concluded that the writing of university-bound students, in the judgement of university instructors, is reasonably acceptable, though many essays do reveal the need for some sort of remedial treatment.

Organization

Effective organization of the essay in the opinion of the scorers and of the teachers reached through the TAI (Table 3.13) was perceived as the most important criterion essays at the interface level should meet. This criterion is also noted under objective 10(a) in the CDQ. A sample of one hundred essays, including the fifty judged by all scorers, was appraised with reference to this criterion (see Appendix A), and the following is the summary of impressions quoted from the Project II Appendix.

Organization: Essays at the bottom of the scale generally exhibited very poor organization, often with poor and infrequent paragraph division. Conclusions, if

reached, tended to be very limp and sometimes essays trailed off altogether.

However, in all the later parts of the range examined, essays, whatever other weaknesses they had, did exhibit adequate to good control of paragraphing and evidence of reasonable planning. Though there was failure in some to give due weight to the more important aspects of their argument, they were reasonably coordinated and developed towards a conclusion.

From the sample examined closely it appeared that essays scored from 4.3 up generally exhibited reasonably effective organization. As a score of 4.6 represents the 10th percentile for Grade Thirteen university-bound students, one may conclude that generally these students do write essays that have from adequate to excellent organization.

Editing

This criterion is also mentioned in association with objective 10(a) in the CDQ. While effective editing of writing involves the clearing out of all errors, error count data for grammar, sentence structure, and diction are examined separately in this report under those headings. Appraisal of "editing" skills, therefore, is confined to examination of error counts for conventions: spelling, capitalization, and punctuation (Table 3.4).

Of the total errors found in the fifty essays examined, errors in these conventions account for more than half: average frequency of convention errors in the sample is 3.8 per 200 running words, while the frequency for all errors in sentence structure, grammar, and diction combined is 3.3.

It is probably on the frequency of error in conventions that writing performance will be judged most severely. Frequency of convention errors above the 75th percentile for university-bound students is 2.2 per 200 running words; for essays between the 75th and 50th percentiles the frequency is 1.9. One may note, albeit with a very limited sampling, that frequency of error in conventions for essays above the 50th percentile is not closely related to score for the essay and that the frequency is rather high, even in the best essays. As noted in the Project II report, it may be that students were not well motivated for the Interface tests and consequently may have been careless in checking their work; nevertheless, the frequency, even in the best writing, of errors in spelling and punctuation deserves some attention.

Diction

Two CDQ objectives related to effective word choice: 12, "Use words with precision", and 13, "Use words with subtlety". The exit-entry expectations for these objectives from Table 3.1 are as follows:

	13 EXIT	UNIV. ENTRY
Use words with precision	4.0	1.5
Use words with subtlety	3.5	1.3

Again the regular discrepancy in use of the 0 - 7 scale occurs, and we have no specific TAI data as a further index of secondary and postsecondary perceptions.

The error count for diction (i.e. words misused, not a measure of less or more effective language choice) as provided in Table 3.4 may be summarized as follows: The frequency of error

for essays above the 75th percentile is 0.25 for 200 running words; for essays between the 50th and 75th percentile, 0.80; and for essays between the 25th and 50th percentile, 0.55. The differences in frequency do not appear to be significant beyond showing that frequency of error is relatively constant across essays above the 25th percentile and so likely has little relationship to the overall evaluation of those essays. It does not appear that wrong word choice is the most serious problem in the writing of essays at the interface.

However, the examination of one hundred essays from the standpoint of effective word choice (see criterion in Appendix A), resulted in the following observation:

Style-Diction: Regrettably, Diederich's description of the dull average, "The writer is addicted to tired old phrases and hackneyed expressions," fairly characterizes the writing in a high proportion of the essays examined. The search for the exact word, with a few notable exceptions, simply was not taking place. It may be speculated that students were insufficiently motivated to put forward their best efforts; nevertheless, the lack of colour and vigour in diction and imagery is discouraging. The most popular topics (1 on Violence and Censorship, and 4 on Course Options--Secondary) particularly appeared to encourage dullness. Pretentiousness of diction was rare (and almost welcome); those students who wanted to show off did so in other ways.

Generally the quality of language choice, with some interesting exceptions, was rather inept in all but those essays of very top rank. Though the sample is relatively small, it would seem that much more needs to be done in writing programs in secondary schools to encourage students to be more vigorous in their search for effective language in their writing.

Argument

Effectiveness of argument was another criterion by which the hundred essays were evaluated, and was also an objective in the CDQ (Table 3.1).

13	UNIV.
EXIT	ENTRY

16. Present an argument
effectively:

3.6	1.5
-----	-----

This was a criterion deemed particularly important by the Project II Advisory Committee in assessing literacy for postsecondary education and a principal reason for the selection of the writing mode for the test. Ability to present a case logically with good supporting evidence and with some objectivity or detachment was considered a particularly important ability for the postsecondary student to have.

Evaluation from this standpoint of the hundred essays produced the following summary:

Argument: The quality and force of the argument with use of appropriate illustrations and giving a balanced treatment proved to be much more unevenly distributed. A few of the otherwise weakest essays did show some strength in this regard, though most of the essays that with some consistency could be praised for quality of argument were found in the portion of the score range next to the top (7.7 - 8.0).

The general weaknesses exhibited included superficiality of generalizations and lack of substance in the way of support. Some writers failed to discriminate between less and more important aspects of the case made, and tended to introduce redundancies or irrelevancies.

Some essays judged weak in most other respects provided quite stimulating and effective argument; essays above 6.8 on the scale were generally competent with respect to argument and evidence. Generally, essays in which students exhibited commitment to and interest in what they were writing, a feature all too infrequent in the essays examined, showed greater effectiveness in argument.

With 6.6 the mean score for essays of university-bound Grade Thirteen students, one can make the generalization that most essays above the mean do reveal capacity to develop an argument effectively, while in essays below the mean the capacity is rather infrequent.

Flavour

Though there is no objective either in the CDQ or TAI which refers specifically to the flavour or colour of the student's writing, the report on essay performance would be incomplete without a summary statement of the assessment of essays in the sample from the standpoint of this criterion. The assessment of the hundred essays produced the following general conclusion:

Flavour: The most popular topics (1 and 4), though sometimes argued strongly, had generally a staleness about them.

"Flavour" was not entirely concentrated in the superior essays. Some middle-range essays, undistinguished in other respects, came alive through the evidence of personal commitment on the part of the writer, while some of the essays in the top of the range, correct in most respects, lacked originality and freshness. While more of the best-rated essays had flavour, it was not a characteristic upon which one could generalize about the range.

The absence of flavour in so many essays again may be a reflection of the testing situation in which motivation was low and students could expect no writer-audience interaction. It would be unwise, therefore, to conclude that most senior secondary students are without the ability to produce flavourful prose.

In fact, the most interesting and enjoyable examples of writing sometimes came from those students who refused to take the test seriously and so entered into a sort of devil-may-care dialogue with the anonymous examiner. The freshness and breeziness of style, while perhaps a little much of a good thing, produced an appealing personal ingredient otherwise generally lacking. That this did occur with some frequency underlines the importance of developing "real" writing situations in order to measure more effectively student writing ability.

In general, as noted earlier for argument, the most effective writing appeared in those essays where the student had commitment and interest. Even when account is taken of the problems created by the testing situation, the general lack of vigour and concreteness in diction, argument and flavour in so many of the essays suggests the need for renewed or increased emphasis on stimulating writing programs in the senior division.

7.2 The Grade Twelve-CAAT Interface

Much of the previous discussion of the writing, especially those sections concerned with the positive characterization of the essays and general impressions of writing quality, have equal pertinence for the Grade Twelve-CAAT interface. Consequently, discussion here will be brief.

According to data previously noted from Tables 3.8.1, 3.8.2, and 3.8.3, teachers of Grade Twelve and CAAT first year instructors see improved writing performance as a very important aim of courses at the interface. Table 3.8.2, however, indicates that the weight given to writing in formal instruction, expressed as a percentage of course time, differs markedly. 49% of Grade Twelve Advanced and 57% of Twelve General teachers state that less than 21% of instructional time is given to writing, while 78% of CAAT teachers claim to concentrate on writing in over 20% of course time. A very substantial 51% claim to give more than 40% of course time to writing.

For the CDQ objective 10(a), "Demonstrate facility in writing...the essay (expository prose)", respective exit-entry expectations on the 0 - 7 scale (Table 3.1) are:

Grade 12A Exit - 4.0
Grade 12G Exit - 3.2
CAAT Entry - 1.3

with the usual discrepancy between CAAT entry and Grade Twelve General exit expectations. According to Table 3.12, 75% of CAAT teachers expect at least 75% of entering students to be able to write an acceptable essay in the mode assigned; 93% of Advanced Twelve and 67% of General Twelve teachers had the same expectation for exit. It appears, therefore, that there is not a great deal of real difference in perception between CAAT and Grade Twelve teachers concerning expectations for Grade Twelve graduates with respect to this writing mode.

The mean score (Table 3.9) for essays of students going directly from Grade Twelve to postsecondary education was 5.3 on the 1 - 10 scale, suggesting that overall performance was at least fair. However, as the scale provides no characterization of the writing, further appraisal of a sample of the writing was conducted according to the methods described earlier in the chapter.

Postsecondary Judgements (Table 3.5)

The classification of the fifty essays as "Acceptable", "Requiring some remedial treatment", and "Unacceptable" was carried out by three instructors from each of the four CAATs. Five of the twelve were members of the English faculty; the remainder were from other divisions.

The problem of sample size is further complicated for the Twelve-CAAT interface by the way in which the randomly selected fifty essays happen to be distributed across the score range. As the right-hand margin of Table 3.5 indicates, there are only three essays that happen to fall between the 25th and 50th percentiles. Generalizations from the data therefore must be very tentative.

Of those essays which in this table fall above the 25th percentile (essays 1 to 34 in rank order), none was judged "Unacceptable" by any CAAT instructor. Of essays between the 90th and 50th percentiles, three were judged in need of remedial treatment by a majority of CAAT instructors, though several more, judged by the response pattern, appeared to be "borderliners". Unfortunately, data for the 25th to 50th percentile range are much too limited for further observations.

The general impression, as far as the data base allows, is that the essays above the 50th percentile were acceptable from the standpoint of CAAT teachers, though likely a good many of the essays in the sample were written by students in Grade Thirteen planning to attend university.

Organization

The summary statement based on this criterion is quoted in the previous section, and applies to the one hundred essays sampled: essays from university-bound, CAAT-bound and workbound Grade Twelve and Thirteen students alike.

As noted there, essays above a score of 4.0 on the holistic scale generally exhibited adequate to good characteristics of organization. As the score for the essays at the 25th percentile for the Twelve-CAAT interface students is 4.4, it is reasonable to conclude that most of the writing of these students does show effectiveness in organization.

Editing (Table 3.4: Error Count for Conventions)

It has already been noted that the frequency of errors in conventions accounts for over half the total errors in the fifty essays examined, and that the frequency is relatively high even in the essays judged best holistically. For those students at the Grade Twelve-CAAT interface, the frequency of errors in spelling, capitalization, and punctuation is 3.8 for 200 running words in those essays between the 50th and 75th percentiles. The number of essays (only three) between the 25th and 50th percentiles renders any further useful observation impossible. This frequency of error in convention appears unacceptably high.

Diction

The combined CDQ objectives (12 and 13) concerning word choice (Table 3.1) provide the following expectations data:

	12A	12G	CAAT
	EXIT	EXIT	ENTRY
Use words with precision	3.6	3.0	1.4
Use words with subtlety	3.0	2.1	1.0

The error count for diction, described earlier, shows (Table 3.4) a frequency of error of 0.7 for those essays falling between the 50th and 75th percentiles. It appears, then, just as for the university-bound Grade Thirteen students, that problem areas in language do not include to a marked extent actual errors in word choice.

On the other hand, student ability to choose words effectively, as ascertained through the appraisal of the hundred essays, leaves much to be desired. The summary impression is given under "Diction" in the preceding section. Generally only the very best essays consistently revealed effectiveness in word choice; very few of those essays (above 7.7 on the holistic scale) would be the product of students at the Grade Twelve-CAAT interface, where the mark at the 90th percentile is 6.7.

Argument

Expectations for the CDQ criterion of argument (Table 3.1) are as follows (scale of 0 - 7):

	12A	12G	CAAT
	EXIT	EXIT	ENTRY
16. Present an argument effectively	3.7	3.3	1.3

Expectation on the part of all teacher groups here is not very high, and results from the appraisal of the hundred essays on this criterion tend to bear out this pessimism. The summary impression, given under "Argument" in the previous section, shows that effectiveness of argument in terms of logic, evidence and balanced treatment, though not confined only to the top essays (scored 7.7 and above), is rather infrequently found below that point, and then chiefly in those welcome essays where the student has shown some vigour and interest in what he is writing about.

Because the score at the 90th percentile for students at the Twelve-CAAT interface is only 6.7, it is reasonable to assume that effectively developed argument is quite infrequent in the writing of these students.

Flavour

Please, in the context of the writing of students planning postsecondary education from Grade Twelve, examine the general

remarks at the conclusion of the preceding section. Perhaps the most serious criticism that can be made of the writing generally was its blandness.

From the standpoint of planning and organization, generally the essays of both groups of students, those proceeding from Grade Twelve to postsecondary and those proceeding from Grade Thirteen to postsecondary, stood up rather well. And in the view of teachers at both levels, this appears to be the criterion that matters most.

While error frequency is unquestionably high, it is important to note for these groups of students that frequency of error in grammar, sentence structure, and diction, considered apart from conventional errors, is actually quite low. The question of what constitutes a tolerable frequency of error is one this study cannot answer, but it is reasonable to say that generalizations such as "Secondary school graduates cannot write a sentence" or "The writing of secondary school graduates is ungrammatical" are untrue, and to suggest that writing competence would be improved by the return of prescriptive grammar teaching would at the least be a misplacement of emphasis.

The error count evidence, on the other hand, does point to a need to encourage students to take a great deal more care in their writing with respect to the conventions of language: punctuation, capitalization (to a lesser extent), and spelling. Many of the errors observed in this sample are evidently a product of carelessness; others suggest that more intensive review of conventions is required.

Even though frequency of error is low for grammar, diction, and sentence structure, more careful assessment of the essays reveals that relatively few students in either group possess adequate competence in style to move beyond very pedestrian (merely "adequate") writing. It has been noted that usually in essays scored below 6.0 holistically there is little evidence that the student consciously uses sentence style to achieve his

effect. With interesting individual exceptions, the quality of word choice exceeds the ordinary in only those essays ranked near the top holistically.

Similarly, effectiveness of argument, though occasional among essays of lower rank, appears with consistency only near the top of the scale.

In general, "flavour", which would reflect a commitment of the writer to what he has to say, and with it the refinement of argument, vigorous language choice, and effective sentence style, are very much lacking in the majority of the essays examined.

Though all data from Project II and Project III indicate that effective writing is and ought to be a major objective of courses in the senior secondary grades and of first year courses in colleges and universities, the evidence both from some of the course description data (Project III Report) and from Table 3.8.2 is that language and writing components of courses in all but the CAATs frequently receive little course emphasis in terms of time allocation.

Generally essays above the 25th percentile did pass the postsecondary appraisal sector. For the universities, none of the essays above the 25th percentile was rejected as unacceptable by a majority of the judges though many between the 25th and 50th percentiles evidently suggest a need for some remedial attention. The same pattern, though the sample available between the 25th and 50th percentile is too small for more than a tentative statement, appears in the CAAT appraisal. No essays with scores above the 25th percentile were found unacceptable, though some required remedial treatment.

The sum of the evidence suggests that the majority of students planning to enter postsecondary institutions from Grade Twelve or from Grade Thirteen do meet a postsecondary "survival" standard in their writing and most are able to organize their writing adequately or well. Given that, there is much more that

could be done to encourage vigour and tightness of writing style and to stimulate sufficient interest on the part of the student to take more care, for the reader's sake, of the conventions of language.

8. CONCLUSION

Chapter Two of the Interproject Analysis Report identifies some of the difficulties faced in the reports on language in relating, through test performance data, specific skills to specific topics or objectives. The kinds of test instruments employed for English do not lend themselves to this type of correspondence. As well, for the reasons elaborated at the beginning of this chapter, it was necessary to restrict evaluation of student performance in English to specifics of reading, language and writing, leaving aside the many objectives related to literature, objectives or topics which evidently compose a high percentage of content in many courses at the interface.

Consequently a review of apparent anomalies, gaps, or duplications must be severely restricted in relation to all possible course objectives and must refer to objectives in rather general terms.

In the interproject analysis of English it will be apparent that no significant duplications have emerged. Expectations generally appear to be either in conformity with the realities of student performance or else overly optimistic. There may indeed be duplications in the area of literature as so high a proportion of time and emphasis is given to this area of English, but this is not discernible from an analysis that has had to be confined to reading, language, and writing.

Presumably, if some undesirable gaps are evident in those areas being examined in this analysis, any redress of imbalances will have its effect in the proportion of course time and

emphasis given to language and literature respectively at one level or another or at all levels involved in the present study. The outcome may be some realignment of priorities affecting expectations with respect to literature.

One "gap" concerning testing in the areas of language, writing and reading has been emphasized at different points in this chapter and should be underlined once again here.

If there are reasons for future testing of students in these areas at the interface, either in senior grades of secondary schools or upon entry to postsecondary institutions, it is evident that the tests employed in this study and likely others in the SACU repertoire are inappropriate and too difficult for all but the university-bound student. They serve no doubt as a screen to separate that student from the rest, but they are not helpful in characterizing performance in language or reading below this level. As well, TAI responses indicate some discomfort with present test format and very strong support for the inclusion of a sample or samples of student writing in any appraisal of students' performance with language.

The various indices used to analyze gaps in expectations at the interface appear to indicate that in most aspects of "literacy" examined secondary school teachers and postsecondary teachers have made a fairly realistic estimate of the competences in English of SSGD and SSHGD students. Though the "preferred" competence level (CDO) and other indicators from the TAI show that all teachers would like to see performance as better than it is, only for reading ability at the Twelve-CAAT interface do we find fairly strong evidence that there is a "gap" in expectation. It appears probably the case that first year CAAT instructors have higher expectations concerning the reading ability of students than present SSGD graduates proceeding to CAATs are able to deliver. It appears that some strategy or strategies should be found -- perhaps more time for reading, perhaps a more formal program of reading instruction--to improve the reading skills of students in general level programs. Note that this "gap" does not

appear to extend to SSHGD students proceeding to university; the relatively greater emphasis placed there on inference skills over reading for literal meaning appears to be in accord with test performance.

There appears also to be a gap, though not as pronounced, at the Grade Twelve-CAAT interface concerning a number of aspects of writing performance, both in general and with reference to specific skills tested in the multiple-choice format. At the CAAT level the problem seems to be appropriately addressed through the substantial weight, in terms of time allocation, given to writing. Perhaps a similar weight should be given writing and language in Grade Twelve programs or sooner, particularly at the general level.

The main gap discerned in interproject analysis is the distance between what teachers in this study say should be emphasized in language and writing and what in fact appears to be the case in terms of time given to writing in a high proportion of courses in all but the CAATs (note again Table 3.8.2.).

It has been noted that serious errors in grammar, word choice, and sentence structure do not appear to characterize the writing of students above the 25th percentile at either interface, and that students generally do show competence in organizing what they have to say.

The two serious concerns in the sample of writing performance examined closely in this study are, first, the frequency of errors in conventions, and, second, the general dullness of the writing.

A shift in course emphasis to provide more writing opportunities would certainly help to close the gap noted above, but something can be said as well about the kind of emphasis needed in writing programs.

There does not appear to be a need for analysis, drills, and reviews concerning the nature of the sentence. Rather the emphasis should be on the creation of more effective stimuli for writing, encouraging the student to write with commitment for a real audience and to experiment with diction and style. A writer with a strong sense of audience is likely to show more respect for the conventions of written English as well. Further, greater emphasis on language and writing in the program will ensure the student has more frequent opportunities to write.

TABLE 3.1

MEAN REPORTED ENTRY AND EXIT LEVELS FOR
OBJECTIVES SELECTED FOR INTERPROJECT ANALYSIS

Levels of competence expected (1) at exit, Grade Twelve
General and Advanced
English, and entry,
CAAT English
(2) at exit, Grade Thirteen
English, and entry,
university English

<u>Objective</u>	<u>Mean reported level of*</u>				
	<u>exit</u>	<u>exit</u>	<u>entry</u>	<u>exit</u>	<u>entry</u>
	<u>12G</u>	<u>12A</u>	<u>CAAT</u>	<u>13</u>	<u>univ.</u>
LANGUAGE					
1. Comprehend a variety of materials (essential meaning and significant details)	3.6	4.2	1.8	4.3	2.1
5. Apply inferential skills	2.8	3.5	1.3	3.9	1.4
8. Write an effective sentence	3.6	4.2	1.9	4.6	1.9
10.(a) Demonstrate facility in writing in terms of planning, organization, presentation and editing essays (expository prose)	3.2	4.0	1.3	4.5	1.8
11. Use language in a grammatically appropriate manner	3.3	3.8	1.8	4.3	1.9

TABLE 3.1 (continued)

<u>Objective</u>	<u>Mean reported level of*</u>				
	<u>exit</u> <u>12G</u>	<u>exit</u> <u>12A</u>	<u>entry</u> <u>CAAT</u>	<u>exit</u> <u>13</u>	<u>entry</u> <u>univ.</u>
12. Use words with precision	3.0	3.6	1.4	4.0	1.5
13. Use words with subtlety	2.1	3.0	1.0	3.5	1.3
16. Present an argument effectively	3.3	3.7	1.3	3.6	1.5

*SCALE EMPLOYED:

- 0 - no competence
- 1 - minimal competence
- 2 -
- 3 - moderate competence
- 4 -
- 5 - competence in varied situations, some originality
- 6 -
- 7 - mastery, competence in high level, creative situations

TABLE 3.2

SUMMARY OF DATA SOURCES EMPLOYED IN INTERPROJECT ANALYSIS

<u>EXPECTATIONS</u>		<u>PERFORMANCE</u>	
Project III Objective (see Table 3.1)	Project II TAI Data	Project II Test Scores	Project II Essay
1	yes	Reading Subtest Scores Mean of raw scores --comprehension items	
5	yes	Reading Subtest Scores Mean of raw scores --inference items	Error Count (Sentence structure)
8	yes	Language Subtest "Construction Shift"	Characterization of essays--Sentence style
11	yes	Language Subtest "Sentence Correction"	Error Count (Grammar)

TABLE 3.2 (continued)

<u>EXPECTATIONS</u>		<u>PERFORMANCE</u>	
Project III Objective (see table 3.1)	Project II IAI Data	Project II Test Scores	Project II Essay
12,13	no		Error Count (Diction)
			Characterization of essays--Diction
16	yes, respecting mode		Characterization of essays--Effectiveness of argument
10(a)	yes	Holistic Scoring of Essays	Postsecondary assessment of acceptability
			Characterization of essays: --Flavour --Organization
			Error Count (Conventions-- editorial skill)

TABLE 3.3
RESPONSE TO ITEMS: READING COMPREHENSION AND LANGUAGE ACHIEVEMENT TEST*

ITEM	FORM 1						FORM 2					
	GRADE 13			GRADE 12			GRADE 13			GRADE 12		
	%R	%NA	Rev %R	%R	%NA	Rev %R	%R	%NA	Rev %R	%R	%NA	Rev %R
	<u>READING</u>											
1	92	0	92	86	0	86	72	1	73	61	1	62
2	77	2	79	61	4	61	82	1	83	73	1	74
3	80	2	82	59	3	59	61	4	64	45	5	47
4	39	4	41	28	4	28	61	3	63	47	4	49
5	15	4	16	8	4	8	27	10	30	21	12	24
6	80	4	83	70	6	70	74	7	80	58	12	66
7	72	8	78	61	12	61	55	11	62	42	14	49
8	57	19	70	53	25	53	51	15	60	35	20	44
9	50	23	65	38	31	38	24	19	30	18	22	23
10	30	19	37	24	26	24	21	30	30	14	31	20

TABLE 3.3 (continued)

ITEM	FORM 1			FORM 2		
	GRADE 13			GRADE 12		
	%R	%NA	Rev %R	%R	%NA	Rev %R
<u>SENTENCE CORRECTION</u>						
			Rev %R	%R	%NA	Rev %R
22	77	0	77	73	0	73
23	81	0	81	64	0	64
24	82	0	82	67	1	68
25	88	0	88	83	0	83
26	40	2	41	34	2	35
27	67	1	68	54	0	54
28	43	1	43	34	0	34
29	76	0	76	67	0	67
30	34	0	34	19	0	19
31	79	1	80	64	1	65
32	54	1	55	37	1	37
33	80	4	83	73	5	77
34	33	3	34	26	2	27
35	70	4	73	66	5	69
36	43	9	47	40	9	44

*%R -Percentage of students obtaining right answer.
%NA -Percentage not responding.
Rev %R -Percentage of students who, having responded to the item, obtained the right answer.

TABLE 3.4

FREQUENCY OF ERRORS IN 50 ESSAYS RANKED BY HOLISTIC SCORING
(ERRORS PER 200 WORDS)

Percentile Divisions: Grade 13- Postsec.	Rank of Essay	Errors in conventions			Errors in sentence, grammar, diction				Sub- total	Sub- total	Total errors	Percentile Divisions: Grade 12- CAAT	
		Sp.	Punc. Cap.		Major	sent. struc.		grammar					diction
						Minor							
75-----	1	0.5	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0		
	2	0.0	1.5	0.5	0.0	1.0	1.5	0.0	0.0	2.5	4.5		
	3	0.5	1.0	0.0	0.0	0.5	1.0	0.0	0.0	1.5	3.0		
	4	3.0	1.5	0.0	0.0	0.5	0.5	0.5	0.5	1.5	6.0		
	5	0.5	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.5		
	6	0.5	2.0	0.0	0.5	0.0	1.0	1.0	1.0	2.5	5.0		
	7	1.0	0.0	0.0	0.0	0.5	0.0	0.5	0.5	1.0	2.0		
	8	0.0	0.5	1.0	0.0	0.0	0.0	0.0	0.0	0.0	1.5		
	9	2.0	2.5	0.0	1.0	0.5	2.0	0.5	0.5	4.0	8.5		
	10	0.0	1.5	0.0	0.5	0.0	0.0	0.0	0.0	0.5	2.0		
50-----	11	0.0	0.0	0.0	0.0	0.5	1.0	0.5	0.5	2.0	2.0		
	11	0.0	2.5	0.0	0.0	0.0	1.0	1.0	1.0	2.0	4.5		
	13	1.5	1.0	0.0	2.0	1.0	3.0	3.5	3.5	9.5	12.0		
	14	0.0	0.5	1.0	0.5	0.5	0.0	0.0	0.0	1.0	2.5		
	15	0.0	0.0	0.0	0.5	0.0	1.0	0.5	0.5	2.0	2.0		
	16	2.0	1.0	1.0	0.0	0.0	0.5	1.5	1.5	2.0	6.0		
	17	0.5	0.0	0.0	0.0	0.0	2.0	1.5	1.5	3.5	4.0		

TABLE 3.4 (continued)

Percentile Divisions: Grade 13- Postsec.	Rank of Essay	Errors in conventions			Sub- total	Errors in sentence, grammar, diction				Sub- total	Total errors	Percentile Divisions: Grade 12- CAAT	
		Sp.	Punc.	Cap.		sent. struc.	Major	Minor	grammar				diction
25-----	18	0.5	2.0	0.0	2.5	0.0	0.0	0.0	0.0	0.0	2.5		
	19	2.5	2.0	0.0	4.5	0.0	2.0	0.0	0.0	0.0	6.5		
	19	1.0	4.0	0.0	5.0	0.0	1.0	1.0	0.5	2.5	7.5		
	21	1.5	1.5	0.0	3.0	1.0	0.5	0.0	1.0	2.5	5.5	75	
	21	0.0	2.0	0.0	2.0	0.5	1.5	0.0	0.0	2.0	4.0		
	23	3.5	3.5	0.0	7.0	1.5	1.5	1.0	0.0	4.0	11.0		
	24	1.5	2.0	0.5	4.0	0.0	0.0	1.5	0.0	1.5	5.5		
	25	4.0	1.5	0.0	5.5	0.0	0.0	0.0	1.0	1.0	6.5		
	26	0.5	1.5	0.0	2.0	1.5	0.5	1.0	1.5	4.5	6.5		
	27	1.0	0.5	0.0	1.5	0.5	0.5	2.0	0.5	3.5	5.0		
10-----	28	4.0	2.5	0.0	6.5	0.0	0.5	1.0	0.5	2.0	8.5		
	29	4.0	2.5	0.0	6.5	4.0	0.5	0.5	0.5	5.5	12.0		
	30	0.5	1.0	0.0	1.5	0.0	0.5	2.0	0.5	3.0	4.5		
	31	2.0	0.5	0.0	2.5	0.0	0.0	0.5	2.0	2.5	5.0	50	
	32	1.0	1.5	0.0	2.5	0.0	0.0	1.0	0.0	1.0	3.5		
	33	1.5	1.5	0.0	3.0	1.5	1.0	0.5	0.0	3.0	6.0		
	34	4.0	--15.0---		*19.0	1.0	0.0	0.0	0.0	1.0	20.9		
	35	1.5	1.5	0.5	3.5	1.0	0.5	4.0	1.0	6.5	10.0		

TABLE 3.4 (continued)

Percentile Divisions: Grade 13- Postsec.	Rank of Essay	Errors in conventions			Sub-total	Errors in sentence, grammar, diction				Sub-total	Total errors	Percentile Divisions: Grade 12- CAAI
		Sp. Punc. Cap.				sent. struc. Major Minor						
						grammar diction						
		Sp.	Punc.	Cap.		Major	Minor	grammar	diction			
36		0.5	1.5	0.0	2.0	1.0	0.0	1.5	1.5	4.0	6.0	
37		3.0	1.5	0.5	5.0	2.5	1.0	0.5	2.0	6.0	11.0	
38		0.5	2.5	0.5	3.0	1.0	0.5	0.5	1.0	3.0	6.0	
39		1.5	2.5	0.0	4.0	1.0	0.0	0.5	0.0	1.5	5.5	
40		1.0	3.0	0.0	4.0	2.0	1.0	1.0	1.0	5.0	9.0	
41		0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	2.0	2.0	
42		4.5	0.5	0.0	5.0	3.0	1.5	1.0	2.0	7.5	12.5	
43		6.5	5.0	0.0	11.5	2.5	0.0	0.0	1.0	3.5	15.0	
44		3.5	0.5	0.0	4.0	1.0	1.0	2.0	1.0	5.0	9.0	
45		5.0	1.0	7.0	13.0	4.0	1.0	5.0	2.5	12.5	25.5	-----10
46		2.5	2.5	1.0	6.0	2.0	3.5	0.5	1.0	7.0	13.0	
47		1.0	2.0	0.0	3.0	0.0	0.0	1.5	0.5	2.0	5.0	
48		5.0	0.5	0.0	5.5	1.5	1.0	7.5	5.0	15.0	20.5	
49		5.5	1.5	1.0	8.0	0.5	0.0	0.0	1.0	1.5	9.5	
50		3.0	1.5	0.0	4.5	2.0	1.0	1.0	1.5	5.5	10.0	

*In this essay, otherwise quite mature in style, the student appeared to be playing with punctuation and capitalization conventions (archy and mehitable style), though this found no favour with error counters who ticked him/her for 19 errors in conventions.

TABLE 3.5

APPRAISAL OF 50 ESSAYS BY POSTSECONDARY INSTRUCTORS

Percentile divisions: Grade 13- Postsec.	Rank of essay	Appraisal by universities*					Appraisal by CAATs*					Percentile divisions: Grade 12- CAAT	
		English faculty		Other faculty			English faculty		Other faculty				
		A	R	X	A	R	A	R	A	R	X		
	1	4	0	0	6	1	1	5	0	0	6	1	0
	2	4	0	0	7	1	0	5	0	0	6	1	0
	3	4	0	0	5	3	0	5	0	0	6	1	0
	4	4	0	0	5	3	0	4	1	0	4	3	0
	5	4	0	0	5	3	0	5	0	0	5	2	0
	6	3	1	0	4	4	0	5	0	0	5	2	0
75-----	7	4	0	0	5	3	0	5	0	0	7	0	0
	8	4	0	0	5	3	0	4	1	0	7	0	0
	9	1	1	2	4	3	1	2	3	0	5	2	0
	10	2	2	0	6	2	0	0	5	0	7	0	0
	11	3	1	0	6	2	0	4	1	0	7	0	0
	11	3	1	0	4	4	0	4	1	0	6	1	0
	13	0	2	2	0	4	4	0	5	0	1	6	0
	14	1	3	0	4	4	0	4	1	0	4	3	0
	15	2	2	0	5	2	1	2	3	0	6	1	0
	16	4	0	0	4	3	1	5	0	0	6	1	0
50-----	17	2	2	0	4	1	3	4	1	0	4	3	0
	18	2	1	1	5	3	0	4	1	0	5	2	0

-----90

TABLE 3.5 (continued)

Percentile divisions: Grade 13- Postsec.	Rank of essay	Appraisal by universities*						Appraisal by CAATs*						Percentile divisions: Grade 12- CAAT
		English faculty			Other faculty			English faculty			Other faculty			
		A	R	X	A	R	X	A	R	X	A	R	X	
36		0	2	2	0	6	2	0	5	0	3	4	0	
37		0	1	3	0	5	3	0	5	0	0	6	1	
38		0	1	3	1	5	2	1	3	1	0	7	0	
39		0	2	2	2	4	2	1	3	1	1	5	1	
40		0	2	2	0	5	3	5	0	0	1	6	0	
41		0	2	2	2	4	2	3	2	0	5	1	1	
42		0	1	3	0	4	4	0	4	1	5	1	1	
43		3	0	1	1	1	6	0	2	3	0	6	1	
44		0	0	4	0	4	4	0	4	1	0	6	1	
45		0	0	4	0	4	4	4	1	0	0	3	4	
46		0	1	3	1	3	4	0	1	4	0	5	2	
47		1	1	2	3	3	2	1	4	0	4	3	0	
48		0	1	3	0	1	7	0	4	1	0	3	4	
49		0	0	4	0	2	6	0	3	2	1	5	1	
50		0	0	4	0	5	3	0	5	0	1	6	0	

*A - acceptable
R - evidence of need for remedial treatment
X - unacceptable

TABLE 3.6.1

TEACHER RESPONSES TO TAI QUESTIONS ON
SUITABILITY OF MODE AND TOPICS IN WRITING TEST

		% <u>13</u>	% <u>12A</u>	% <u>12G</u>	% <u>CAAT</u>	% <u>Univ.</u>
Are there other modes of writing which from the standpoint of general literacy are as important as, or more important than, the mode being examined?	NO	50	53	58	48	54
Was the present assignment at a reasonable level of difficulty for students in courses at this level?	YES	94	99	86	94	100
Was the restriction to a single mode fair to students in a test of writing competence?	YES	81	78	71	62	92
Given the restriction to a single mode, what is your opinion of the range of topics?						
	GOOD	58	55	47	38	40
	SATISFACTORY	36	40	40	53	48
	UNSATISFACTORY	6	5	13	9	12

TABLE 3.6.2

TEACHER RESPONSES TO TAI QUESTIONS ON
APPROPRIATENESS OF THE LANGUAGE SUBTESTS

(1) CONTENT

What is your evaluation of the language achievement items in the tests from the standpoint of emphasis given to usage, style, grammar, structure, and idiom?

	% 13	% 12A	% 12G	% CAAT	% UNIV
They provide a reasonable balance in testing important areas of language achievement.	63	63	70	81	71
Though testing a number of important areas, there is an imbalance in emphasis.	17	12	15	6	4
Important areas of language achievement are omitted or tested too lightly.	19	24	14	13	25

Those giving either of the last two responses concerning balance and emphasis referred less to these matters than to areas of language competence they felt should also have been tested. To the extent that omissions were specified, we found approximately a dozen identified, no particular one with high frequency. The most common complaint, and it was infrequent, was that more obvious or crude errors (e.g., sentence fragments, subject-verb agreement) went untested in favour of more subtle and, presumably, less urgent matters. Very few criticized omission of the testing of mechanics and such conventions as punctuation, capitalization, or spelling, even at the General 12 level.

TABLE 3.6.2 (continued)

(2) LEVEL OF DIFFICULTY

What is your evaluation of the "construction shift" items?	% 13	% 12A	% 12G	% CAAT	% UNIV
Suitable but too easy	8	9	2	0	0
Suitable and right for difficulty	52	49	33	38	60
Suitable but too difficult	10	16	42	24	8
Unsuitable for reasons other than difficulty	29	27	23	38	32
What is your evaluation of the "sentence correction" items?	% 13	% 12A	% 12G	% CAAT	% UNIV
Suitable but too easy	13	9	2	0	4
Suitable but right for difficulty	67	72	47	76	86
Suitable but too difficult	10	15	38	18	0
Unsuitable for reasons other than difficulty	9	4	12	6	9
What is your overall evaluation of the difficulty level of the language achievement parts of the tests?	% 13	% 12A	% 12G	% CAAT	% UNIV
Too easy	3	1	0	0	9
Somewhat easy	11	7	0	3	13
About right	53	51	26	38	61
Somewhat difficult	29	31	39	53	17
Too difficult	5	10	36	6	0
	--	--	--	--	--
Mean	3.21	3.43	4.10	3.71	2.86

TABLE 3.6.2 (continued)

(3) COMPARATIVE APPROPRIATENESS OF LANGUAGE TEST AND WRITING TEST

Do you think a test in multiple-choice format is in general a satisfactory means of assessing students' language competence?

%	%	%	%	%
13	12A	12G	CAAT	UNIV

yes	28	30	36	24	12
-----	----	----	----	----	----

yes (qualified)	44	45	37	61	76
-----------------	----	----	----	----	----

no	28	25	27	15	12
----	----	----	----	----	----

Where multiple-choice tests are used, should they be supplemented by other measures?

%	%	%	%	%
13	12A	12G	CAAT	UNIV

yes	87	90	86	82	96
-----	----	----	----	----	----

What importance do you place upon a sample of the student's writing in an evaluation of language competence?

%	%	%	%	%
13	12A	12G	CAAT	UNIV

Essential	93	89	79	91	96
-----------	----	----	----	----	----

Important but not essential	7	11	18	9	4
-----------------------------	---	----	----	---	---

Of minimal importance and utility	1	0	2	0	0
-----------------------------------	---	---	---	---	---

Neither important nor useful	0	0	1	0	0
------------------------------	---	---	---	---	---

TABLE 3.6.2 (continued)

How do you regard the use of both a multiple-choice test of language achievement and a sample of writing in assessing language competence?

%	%	%	%	%
13	12A	12G	CAAT	UNIV

The multiple-choice test is satisfactory by itself.

2	1	3	5	0
---	---	---	---	---

The use of both is important.

69	73	68	80	84
----	----	----	----	----

The writing sample is satisfactory by itself.

28	23	26	15	16
----	----	----	----	----

Neither is particularly satisfactory.

1	3	3	0	0
---	---	---	---	---

If a student's relative standing on a writing test were different from the student's relative standing on a multiple-choice test of the type administered in this study, which would you consider the most valid measure of the student's language competence?

%	%	%	%	%
13	12A	12G	CAAT	UNIV

The score on the writing test

49	53	56	44	56
----	----	----	----	----

The score on the multiple-choice test

3	0	2	0	0
---	---	---	---	---

A combined score weighted in favour of the writing test

42	43	39	38	40
----	----	----	----	----

TABLE 3.6.2 (continued)

A combined score weighted in favour of the multiple- choice test	3	2	0	3	0
A combined score giving equal weight to both tests	3	3	3	15	4

A significant percentage responding to these items strongly objected to the use of multiple-choice tests at all as a measure of language achievement, the higher proportion of strong negatives being found at the secondary panel.

TABLE 3.6.3

TEACHER RESPONSES TO TAI QUESTIONS ON
DIFFICULTY AND APPROPRIATENESS OF READING SUBTESTS

Are the four passages representative of the material you would expect students at this level to be able to read with comprehension?	% 13	% 12A	% 12G	% CAAT	% UNIV
Yes	72	50	14	52	80
What is your overall assessment of the difficulty level of the four passages given in the two forms of the test?	% 13	% 12A	% 12G	% CAAT	% UNIV
Too easy	1	0	1	3	0
Somewhat easy	5	3	1	3	0
About right	45	32	11	26	71
Somewhat difficult	42	40	36	47	29
Too difficult	8	25	51	21	0
	—	—	—	—	—
Mean	3.50	3.86	4.35	3.80	3.29

Only the secondary school teachers were asked to respond concerning the specific passages:

Assessment of difficulty level of each passage (1 (too easy) to 5 (too difficult)) reported as means:

	13	12A	12G
Passage on Mass Marketing	3.14	3.59	4.10
Passage on Viviparity	3.39	3.72	4.17
Passage on Object-Perception	3.81	4.23	4.39
Passage on Mackenzie King	3.27	3.61	4.20

TABLE 3.6.3 (continued)

Is the passage appropriate except for difficulty?		% 13	% 12A	% 12G
Mass Marketing	Yes	82	81	73
Viviparity	Yes	58	56	61
Object-Perception	Yes	58	56	54
Mackenzie King	Yes	87	80	74

TABLE 3.7

ANALYSIS OF OBJECTIVES TESTED: MULTIPLE-CHOICE TESTS

(1) FIRST LANGUAGE SUBTEST--"CONSTRUCTION SHIFT":
ITEMS 11-21, FORMS 1 AND 2 COMBINED

Types of Items and Numbers of Each Type

Item type	Number of items
Change in voice, tense, mood of verb	5
Change of verb to noun	3
Change in clause structure	7
Change in idiomatic structure	1
Miscellaneous changes (e.g., adjective to adverb, negative to positive, interrogative to assertive, change of subject)	6
Total	22

(From V.R. D'Oyley: An updated technical note on CTEL, 1975)(2) SECOND LANGUAGE SUBTEST--"SENTENCE CORRECTION":
ITEMS 22-36, FORMS 1 AND 2 COMBINED*

Idiom - 1)	
Indefinite "it" - 1) Stylistic	5
Redundancy - 3)	
<u>Verbs:</u>)	
Time sequence - 4-1/2)	6-1/2
S-V agreement - 2)	
Coordination,)	
parallelism problems)	3-1/2
Verbals		3-1/2

TABLE 3.7 (continued)

Pronoun references:)	
consistency - 3-1/2)	4
nominative - 1/2)	
No error (i.e., the original)	
was correct--option A in)	6
all items))	
Miscellaneous		1-1/2

*1/2 means stem (plus first option) has two errors, so "half" is assigned to each appropriate heading.

"Style" is a little better represented than appears from the classification, for among distracters the student has to choose in some instances among less and more effective manners of expression.

COMMENT

In some items the key is less than impeccable standard written English. It is presumably awkward to a degree that makes it less distinguishable from the distractors. If this is a necessary step in making the items "work", it is an unfortunate one. The student is forced to pick the best of a bad lot rather than to discriminate excellence or, in a few instances, even a satisfactory level of "standard written English".

(3) READING COMPREHENSION SUBTEST: ITEMS 1-10, FORMS 1 AND 2 COMBINED

(Memorandum to Dr. Alan King from Peter Evans, dated March 22, 1976)

English Testing and Project 3 Matrix

READING COMPREHENSION

(A follow-up to my memo of March 16: I - Reading Comprehension)

For several practical reasons connected with test administration, I have had to change the reading test from CSAT 1970, Form 131, to CSAT 1973, Form 543.

TABLE 3.7 (continued)

- (1) My content analysis of the items does not change significantly the line-up between your matrix and what we are testing, except, I am pleased to say, that there is a now slightly larger proportion of items dealing with inference and implication.

Total literal meaning:	12
Main idea (implied):	2
Inference, implication:	5
Other:	1

As before, the caution remains that this summary is on my examination and interpretation only.

The relationship of testing to matrix and my recommendations (see p. 4 of March 16 memo) remain as they were.

- (2) The kinds of reading material that appear in CSAT 1973:

Mass culture and advertising--
expository, sociological emphasis

Sexual and a-sexual reproduction
among insects--biological "text
book" style

Perception and movement--psychology,
moderately technical

Church, society, economics--
argumentative in tone, historical
in context

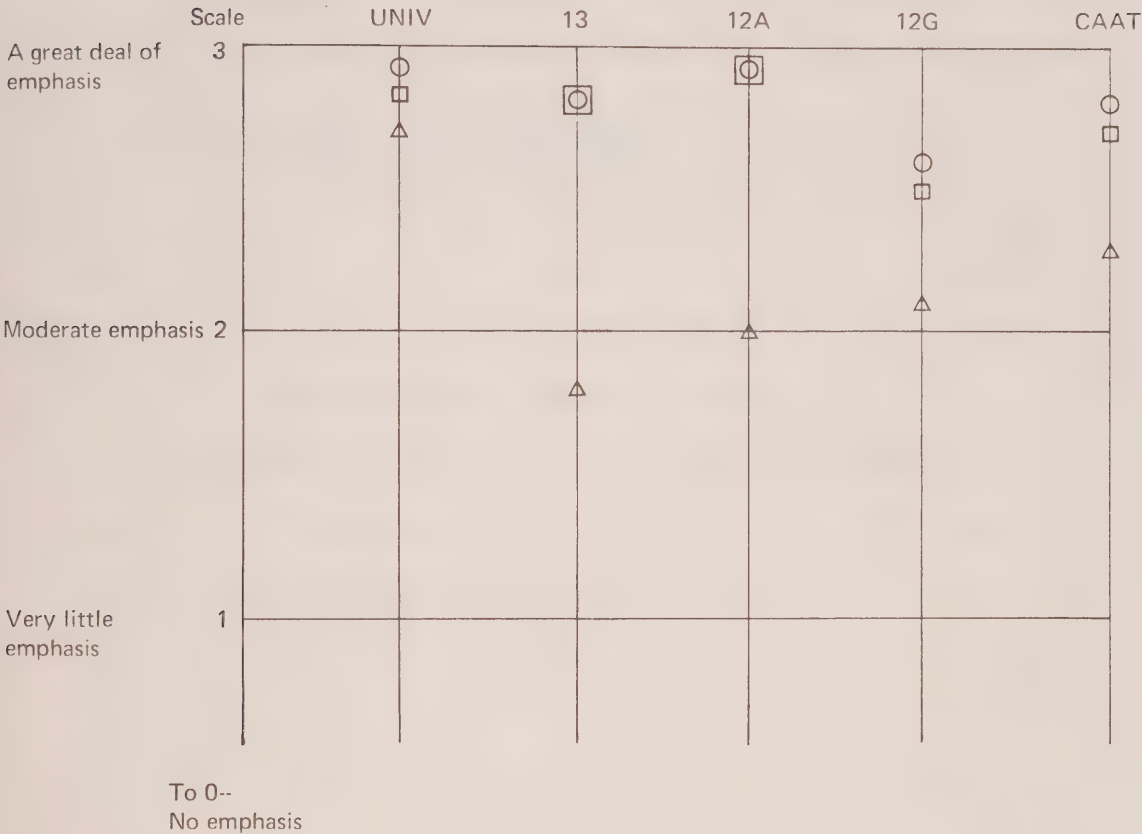
The range of styles thus is a little narrower than in CSAT 1970 but generally adequate in that all are kinds of writing that one would expect a student heading for postsecondary education to be able to read intelligently.

TABLE 3.8.1

RELATIVE EMPHASIS GIVEN THREE AIMS

(COURSE DESCRIPTION QUESTIONNAIRE: "AIMS OF COURSE")

(SCALE 0 - 3; MEAN SCORES)



Code: ○ Promote fluent and grammatically acceptable written English

△ Improve student's reading ability

□ Develop student's ability to organize and integrate ideass and materials

TABLE 3.8.2

TEACHER RESPONSES TO CDQ QUESTIONS ON
PERCENTAGE OF FORMAL INSTRUCTION TIME
GIVEN TO LANGUAGE SKILLS

(A) GRADE 13 AND UNIVERSITY TEACHERS

FORMAL INSTRUCTION TIME GIVEN TO LANGUAGE SKILLS
(READING, WRITING, SPEAKING, ETC. COMBINED)

% of course time	% of instructors responding	
	<u>Grade 13</u>	<u>University</u>
0%-10%	4	50
11%-20%	22	15
21%-40%	66	27
41% or more	8	8

(B) GRADE 12 AND CAAT TEACHERS

FORMAL INSTRUCTION TIME GIVEN TO READING

% of course time	% of instructors responding		
	<u>Grade 12A</u>	<u>Grade 12G</u>	<u>CAAT</u>
0%-10%	78	56	46
11%-20%	16	26	28
21%-40%	6	12	14
41% or more	0	6	12

(C) GRADE 12 AND CAAT TEACHERS

FORMAL INSTRUCTION TIME GIVEN TO WRITING

% of course time	% of instructors responding		
	<u>Grade 12A</u>	<u>Grade 12G</u>	<u>CAAT</u>
0%-10%	10	16	20
11%-20%	39	41	2
21%-40%	49	41	27
41% or more	2	2	51

TABLE 3.8.3

PERCENTAGE OF INSTRUCTORS REPORTING THAT THEY
GIVE "HEAVY" OR "MODERATELY HEAVY" EMPHASIS TO
READING, LANGUAGE AND WRITING OBJECTIVES
(TAI DATA)

<u>Objective</u>	<u>UNIV</u>	<u>13</u>	<u>12A</u>	<u>12G</u>	<u>CAAT</u>
Reading : Literal understanding	56	61	68	81	75
: Main idea	92	90	89	94	84
: Inference	92	94	92	57	69
<hr/>					
Language : Competence in skills tested on language achievement tests	42	59	62	55	77
<hr/>					
Writing : (In the mode assigned)	96	90	97	78	87
<hr/>					

TABLE 3.9

STUDENT PERFORMANCE ON ENGLISH TESTS:
 GRADE TWELVE-POSTSECONDARY AND
 GRADE THIRTEEN-POSTSECONDARY

	10th	25th	PERCENTILE		
			<u>50th</u>	75th	90th
(A) TOTAL TEST SCORE (Scale: -9 to 36)*					
GRADE 13-POSTSECONDARY	7.8	12.4	<u>17.2</u>	22.6	26.5
GRADE 12-POSTSECONDARY	3.0	5.8	<u>8.8</u>	13.2	19.0
(B) READING COMPREHENSION SUBTEST (Scale: -2.75 to 11)*					
GRADE 13-POSTSECONDARY	1.6	3.2	<u>5.1</u>	7.2	8.3
GRADE 12-POSTSECONDARY	---	0.6	<u>2.3</u>	4.0	6.0
(C) LANGUAGE ACHIEVEMENT SUBTEST "CONSTRUCTION SHIFT" (Scale: -2.75 to 11)*					
GRADE 13-POSTSECONDARY	0.5	2.3	<u>4.3</u>	6.5	8.3
GRADE 12-POSTSECONDARY	---	0.5	<u>2.2</u>	3.9	5.8
(D) LANGUAGE ACHIEVEMENT SUBTEST "SENTENCE CORRECTION" (Scale: -3.75 to 15)*					
GRADE 13-POSTSECONDARY	3.4	5.9	<u>8.0</u>	10.2	12.1
GRADE 12-POSTSECONDARY	1.7	3.2	<u>5.5</u>	7.5	9.4
(E) ESSAYS: HOLISTIC SCORING (Scale: 1 to 10)					
GRADE 13-POSTSECONDARY	4.6	5.7	<u>6.6</u>	7.4	8.1
GRADE 12-POSTSECONDARY	3.6	4.4	<u>5.3</u>	6.1	6.7

*The apparent oddity of the scale is owing to the inclusion of the correction for guessing and the equating of the two forms of the test.

TABLE 3.10

TEST APPRAISAL INVENTORY RESPONSES: ENTRY EXPECTATIONS:
 READING FOR LITERAL UNDERSTANDING AND INFERENCE

How many of the students <u>entering</u> courses at this level should have the ability to read a passage for <u>literal understanding</u> ?	% 13	% 12A	% 12G	% CAAT	% UNIV
All the students	73	61	47	72	83
More than 75% of the students, but not all	19	18	19	19	17
	—	—	—	—	—
Total	92	79	66	91	100
How many students <u>on entry</u> should be able to <u>draw inferences</u> and <u>see implications</u> ?	% 13	% 12A	% 12G	% CAAT	% UNIV
All the students	27	15	8	31	42
More than 75% of the students, but not all	37	20	20	19	38
	—	—	—	—	—
Total	64	35	28	50	80

TABLE 3.11

TEST APPRAISAL INVENTORY RESPONSES:
ENTRY-EXIT EXPECTATIONS:
LANGUAGE SKILLS TESTED

How many students should have the competences assessed by the language achievement items upon entry to your course?

(Postsecondary: "institution" replaces "course")	% 13	% 12A	% 12G	% CAAT	% UNIV
All	36	26	23	41	56
More than 75%, but not all	34	23	15	28	28
51% to 75%	24	34	28	13	8
26% to 50%	2	11	20	9	4
One or more, but less than 26%	1	3	10	3	4
None	3	3	3	6	0

How many of the students who successfully complete English courses at this level should have the competences assessed by the language achievement items?

	% 13	% 12A	% 12G	% CAAT	% UNIV
All	44	37	17	28	70
More than 75%, but not all	39	35	28	31	13
51% to 75%	12	18	31	24	9
26% to 50%	1	4	13	10	0
One or more, but less than 26%	1	0	9	0	4
None	3	5	3	7	4

TABLE 3.12

TEST APPRAISAL INVENTORY RESPONSES:
ENTRY-EXIT EXPECTATIONS:
THE ESSAY

How many students should be able to write an acceptable essay of this type upon entry to English courses at this level? (Postsecondary inventories phrased it as "entry to your postsecondary institutions".)

	% 13	% 12A	% 12G	% CAAT	% UNIV
All	66	35	24	45	80
More than 75% but not all	23	38	16	30	16
51% to 75%	9	23	33	15	4
26% to 50%	1	3	20	10	0
One or more, but less than 26%	1	1	7	0	0
None	0	0	0	0	0

How many of the students who successfully complete English courses at this level should be able to write an acceptable essay of the type required in the test? (For postsecondary: "the first year English course you teach")

	% 13	% 12A	% 12G	% CAAT	% UNIV
All	77	60	23	47	88
More than 75% but not all	19	33	44	37	12
51% to 75%	3	6	20	16	0
26% to 50%	1	1	11	0	0
One or more, but less than 26%	0	0	1	0	0
None	0	0	0	0	0

TABLE 3.13

TEST APPRAISAL INVENTORY RESPONSES:
RELATIVE WEIGHT GIVEN CRITERIA FOR EVALUATION OF THE ESSAY

(Mean scores on a scale of 1 (high) to 5 (low).)

	13	12A	12G	CAAT	UNIV
Organization	1.66	1.45	1.68	1.76	1.48
Logic; use of evidence	2.11	2.14	2.23	2.21	1.76
Style (chiefly the sentence)	2.77	2.87	3.05	3.55	2.72
Grammar, usage, mechanics	3.10	2.94	2.95	2.94	2.36
Diction	3.70	3.71	3.82	3.88	3.36

APPENDIX 3A

CHARACTERIZING THE ESSAY

The criteria applied were in part derived from Diederich (1976), pp. 55-57.

1. ORGANIZATION (from Diederich, unmodified)

High: The paper starts at a good point, has a sense of movement, gets somewhere, and then stops. The paper has an underlying plan that the reader can follow; he is never in doubt as to where he is or where he is going. Sometimes there is a little twist near the end that makes the paper come out in a way that the reader does not expect, but it seems quite logical. Main points are treated at greatest length or with greatest emphasis, others in proportion to their importance.

Middle: The organization of this paper is standard and conventional. There is usually a one-paragraph introduction, three main points each treated in one paragraph, and a conclusion that often seems tacked on or forced. Some trivial points are treated in greater detail than important points, and there is usually some dead wood that might better be cut out.

Low: This paper starts anywhere and never gets anywhere. The main points are not clearly separated from one another, and they come in a random order--as though the student had not given any thought to what he intended to say before he started to write. The paper seems to start in one direction, then another, then another, until the reader is lost.

2. EFFECTIVENESS OF ARGUMENT OR VIEWPOINT

High: The writer provides a balanced argument or viewpoint, treating the subject with an appropriate detachment and some originality, and the argument develops to an appropriate conclusion based on the evidence. Judicious use is made of examples and other forms of evidence, and generalizations are supported. The argument of viewpoint develops logically. The writer reveals commitment to his position.

Middle: The writer tends to rely too much on unsubstantiated generalizations and general clichés of thought. The viewpoint expressed may be rather one-sided without adequate regard for alternative positions. Transitions between stages of the argument may not always be clear. The student may evince little commitment to what he is saying and betray a low interest in the subject.

Low: Generalizations and thought clichés abound with little regard for substantive evidence. The argument may be heavily biased or, perhaps worse, the writer may attempt to present all sides without ever taking a firm position and the essay rambles to no conclusion in particular. There is little awareness of logical development and the arguments themselves are tired.

3. STYLE (SENTENCE STRUCTURE)

High: Sentences are well varied in relation to purpose. Less important ideas are effectively subordinated and there is clear evidence that such devices as parallelism are used to good effect. Sentences are varied in length and made economical in expression through ellipsis and tautness of phrase.

Middle: Though some variety in sentence structure is evident, the variety is not particularly well related to effect. The writer appears unaware of the effective use of stylistic and rhetorical features of language. There is limpness of style with occasional minor errors.

Low: No awareness of style or variety in sentence structure is evident, and errors, major and minor, do occur. Sentences tend to ramble to no purpose and expression is woolly or redundant.

4. STYLE (DICTION)

High: The writer uses a sprinkling of uncommon words or of familiar words in an uncommon setting. He shows an interest in words and in putting them together in slightly unusual ways. Some of his experiments with words may not quite come off, but this is such a promising trait in a young writer that a few mistakes may be forgiven. For the most part, he uses words correctly, but he also uses them with imagination.

Middle: The writer is addicted to tired old phrases and hackneyed expressions. If you left a blank in one of his sentences, almost anyone could guess what word he would use at that point. He does not stop

to think how to say something; he just says it in the same way as everyone else. A writer may also get a middle rating on this quality if he overdoes his experiments with uncommon words; if he always uses a big word when a little word would serve his purpose better.

Low: The writer uses words so carelessly and inexactly that he gets far too many wrong. These are not intentional experiments with words in which failure may be forgiven; they represent groping for words and using them without regard to their fitness. A paper written in a childish vocabulary may also get a low rating on this quality, even if no word is clearly wrong.

5. FLAVOUR (from Diederich, unmodified)

High: The writing sounds like a person, not a committee. The writer seems quite sincere and candid, and he writes about something he knows, often from personal experience. You could not mistake this writing for the writing of anyone else. Although the writer may assume different roles in different papers, he does not put on airs. He is brave enough to reveal himself just as he is.

Middle: The writer usually tries to appear better or wiser than he really is. He tends to write lofty sentiments and broad generalities. He does not put in the little homely details that show that he knows what he is talking about. His writing tries to sound impressive. Sometimes it is impersonal and correct but colourless, without personal feeling or imagination.

Low: The writer reveals himself well enough but without meaning to. His thoughts and feelings are those of an uneducated person who does not realize how bad they sound. His way of expressing himself differs from standard English, but is not his personal style; it is the way uneducated people talk in his neighbourhood. Sometimes the unconscious revelation is so touching that we are tempted to rate it high on flavour, but it deserves a high rating only if the effect is intended.

APPENDIX 3B

SELECTED BIBLIOGRAPHY

(1) A SELECTED BIBLIOGRAPHY: ENGLISH TESTING

- Cholvat, J.A. "An examination of Grade 13 marks and OACU test battery scores as predictors of grades obtained in first year Engineering. Informal report, North York Board of Education, n.d. (ca. 1968).
- D'Oyley, V.R. and Gerhardt, G. Technical report: DECAT-0A. Toronto: The Ontario Institute for Studies in Education, 1967.
- D'Oyley, V.R. "Evaluation of English," The Bulletin, May, 1968, pp. 164-165, 185.
- D'Oyley, V.R. and Scott, D.M. "Evaluation of English". Paper read at C.C.T.E. Convention, August 1968 (subsequently revised and published in The English Quarterly under the title "Objective testing and the evaluation of English: some comments").
- D'Oyley, V.R. and Gerhardt, G. "The measurement of English achievement". Test Development Papers No. 2, Paper 3. Toronto: The Ontario Institute for Studies in Education, 1967.
- D'Oyley, V.R. An updated technical note on "The Canadian Test of English Language" (CTEL), unrevised edition. Council of Ontario Universities, Dec. 1975.
- Godshalk, F., Swineford, F. and Coffman, W.E. The measurement of writing ability. College Entrance Examination Board, 1966.
- Khan, S.B. "Validation studies of the Ontario Tests for Admission to College and University". OACU Research Report No. 1. Toronto: The Ontario Institute for Studies in Education, 1967.
- Mellon, J.C. National Assessment and the teaching of English. National Council of Teachers of English, 1975.
- Task Force on Measurement and Evaluation in the Study of English. Common sense and testing in English. National Council of Teachers of English, 1975.

(2) A SELECTED BIBLIOGRAPHY: WRITING EVALUATION

Anderson, C.C. "The new STEP essay test as a measure of composition ability," Educational and Psychological Measurement XX, Spring 1960, pp. 95-102.

Braddock, R., Lloyd-Jones, R. and Schoer, L. Research in written composition. Champaign, Illinois: National Council of Teachers of English, 1963.

Britton, J.N., Martin, N.C. and Rosen, H. Multiple marking of English compositions. London: Her Majesty's Stationery Office, 1966.

Cast, B.M.D. "The efficiency of different methods of marking English compositions," British Journal of Educational Psychology IX (3), 1939, pp. 257-269 and X (1), 1940, pp. 49-60.

Chase, C.I. "The impact of some obvious variables on essay test scores," Journal of Educational Measurement 5, 1966, pp. 315-318.

Coffman, W.E. and Kurfman, D.A. "A comparison of two methods of reading essay examinations," American Educational Research Journal 5, 1968, pp. 99-107.

Cooper, C.R. "Measuring growth in writing," English Journal 64 (3), March 1975.

Diederich, P.B. Measuring growth in English. Urbana, Illinois: National Council of Teachers of English, 1974. This contains a summary of a now out-of-print work: Diederich, P.B., French, J.W. and Carlton, S.T. Factors in judgments of writing ability, Research Bulletin RB-6-15. Princeton: Educational Testing Service, 1961.

Finlayson, D.S. "The reliability of the marking of essays," British Journal of Educational Psychology XXI (2), 1951, pp. 126-134.

Follman, J.C. and Anderson, J.A. "An investigation of the reliability of five procedures for grading English themes," Research in the Teaching of English 1 (2), 1967, pp. 190-200.

Godshalk, F.I., Swineford, F., and Coffman, W.E. The measurement of writing ability. College Entrance Examination Board, New York, 1966.

Mellon, J.C. National Assessment and the teaching of English. Urbana, Illinois: National Council of Teachers of English, 1975.

Myers, A.E., Coffman, W.E., and McConville, C.B. "Simplex structure in the grading of essay tests," Educational and Psychological Measurement 26 (1), 1966, pp. 41-54.

National Assessment of Educational Progress, Report No. 8: Writing: National Results--Writing Mechanics. Washington, D.C.: U.S. Government Printing Office, 1972.

Stalnaker, J. "The essay type of examination," in Educational Measurement, ed. by E.L. Lindquist. Washington, D.C.: American Council of Education, 1951.

Wiseman, S. "The marking of English composition in grammar school selection," British Journal of Educational Psychology XIX (3), 1949, pp. 200-209.

Wiseman, S. "Reliability and validity," British Journal of Educational Psychology XXVI, Nov. 1956, pp. 172-179.

Wiseman, S. and Wrigley, J. "Essay reliability: the effect of choice of essay title," Educational and Psychological Measurement 18 (1), 1958.

(3) UNIVERSITY STUDY OF ENGLISH PERFORMANCE

Norman, C. The Queen's English: standards of literacy among undergraduates in the Faculty of Arts and Science at Queen's University, 1975-76. Department of English, Queen's University, 1976.

CHAPTER FOUR

MATHEMATICS: THE SECONDARY SCHOOL-UNIVERSITY INTERFACE*

1. INTRODUCTION

The official statement of the program of studies in mathematics for Grade Thirteen students in Ontario in 1976 is given in the document "Mathematics: Ministry of Education, Ontario, Senior Division, 1972". A description of the actual school programs is given in detail in the Project III report (King et al., 1976). There is a close correspondence between the specifications of the Ministry document and the content of these courses.

A brief look at some of the events in mathematics education at this level will show that the official program has had time to become well established in Ontario schools, and that the standardized achievement test chosen for use in this study has proven useful in the past.

Approximately twenty years ago, there was a climate of general dissatisfaction with the then current secondary school mathematics program, and with the degree of integration of this program with university programs. The feeling of the general public at that time was in many ways comparable to the present public uneasiness. Mathematics educators at both secondary and postsecondary levels shared this general concern, and took initiatives toward finding solutions for the apparent problems.

*(This chapter was prepared with the assistance of John Turner, who teaches Mathematics at Centennial Secondary School in Belleville.)

One result of these initiatives was the formation of the Ontario Mathematics Commission, a coalition of all interested parties. It was charged with the responsibility of reorganizing the mathematics programs from Kindergarten to Grade Thirteen in such a way that not only would serious gaps and duplications be removed, but also more mathematical substance would be contained in the course at all levels.

After intensive study by this Commission, new programs were developed for all grades. They were introduced throughout the province during the period from 1960 to 1970. During the late 1960s, a team at the Ontario Institute for Studies in Education, headed by Dr. D.M. Horn, developed the Ontario Mathematics Achievement Test. It was specifically tailored to the new programs which were being adopted by the Department of Education, and the team relied heavily on advisers from the Ontario Mathematics Commission. There is reason, then, to expect a good match between this test and the school program, within the kinds of limitations which are imposed by test standardization (see Chapter 2).

1.1 Difficulty Indices for Standardized Tests

It is important, when considering student performance on a test as evidence bearing on the existence of program duplication or gaps, to bear in mind the level of difficulty of the test items. On this particular test, as on many standardized tests, the typical item is one which is answered correctly by between 50% and 60% of the students. Items of this difficulty are well suited to provide the kind of discrimination among students which is the purpose of the test.

One could arbitrarily define a "gap" in such a way that, on a typical item answered correctly by between 50% and 60% of students, the unsuccessful balance of 40% to 50% were said to fall into this gap, if the item content is material they are assumed to know when they enter university. This definition,

however, does not take into account the fact that the difficulty of a test item can be (and is) manipulated through different choices of words or numbers, or in other ways, so that even if all the students writing the test have learned the material it is expected that only about half of them will be successful on the item.

It is equally defensible, then, to propose that if one-half of the candidates correctly answer the item, then it can be assumed that all of them know the material, and that there is no gap. Using this definition, one might say that if the success rate is, for example, 20%, then a gap exists, since this result is 30% below the accepted standard.

Since both these definitions are highly arbitrary, and since each is misleading in a different direction, we propose to report estimates of the magnitude of gaps and duplications with what may appear to be excessive caution.

1.2 Correspondence of Test to Program

The Project II report describes the process of mathematics test selection, and the Project III report the process of construction of the content matrix for mathematics. It is sufficient here to state that the working team for each project was composed of university teachers, community college teachers, and secondary school teachers of both Grade Twelve and Grade Thirteen mathematics, along with members of the respective project staffs. As an aid to interproject coordination and eventual interproject analysis, the teams had four common members. This overlap was helpful in effecting some integration of the projects, but did not distort the essentially different basic purposes of the teams.

The goal of the Project III team was the designing of instruments which could be used to provide a detailed description of the content of the mathematics programs actually in use in Ontario secondary schools at the Grade Twelve and Grade Thirteen

levels and in community colleges and universities in their introductory year, by eliciting responses from teachers at all levels. The Project II team was to select the best available test at each of the Grade Twelve and Grade Thirteen levels to measure the achievement of students in mathematics. It seemed desirable, if not essential, to arrange the teams' work schedules so that the construction of the Project III instruments would not be influenced by Project II's choice of tests. It was also necessary to employ safeguards in the gathering of data from teachers, so that responses to the Project III instruments would not be biased by teachers' knowledge of which topics had appeared in the tests.

At the Grade Thirteen level, the Project II team selected as the best available test one form of the Ontario Mathematics Achievement Test, renamed for its use in the study the Mathematics Achievement Test, and translated into French for use in Francophone schools as the Test de rendement en mathématiques. At the same level, the Project III team produced a Course Description Questionnaire, which included a content matrix listing all those topics which the team felt it was at all likely would be dealt with in Grade Thirteen or first year university mathematics courses. Teachers at both levels were asked to give information about student performance on each of these topics. A list of the topics included in the matrix under the headings "Relations and Functions" and "Calculus" may be found in Appendix 4A. Only this portion of the matrix is included here, since the content of the test bears only on these areas of mathematics, and the students tested were chosen only from among those enrolled in both Relations and Functions and Calculus.

Appendix 4B is a diagram showing the relationship among the content of, respectively, the Mathematics Achievement Test, the Ministry of Education guidelines, and the portion of the content matrix included in Appendix 4A. The lowest circle represents the domain of the test, the next the domain of the guidelines, and the highest the domain of this portion of the matrix. The vertical line somewhat arbitrarily separates the parts of those

domains dealing with calculus from the parts dealing with relations and functions.

A comparison of the content matrix and the Ministry guidelines shows that space E is empty. The guidelines cover all topics listed under "Relations and Functions" in the matrix. Space F contains a very large number of topics--not surprisingly, since the great majority of topics in calculus are generally not dealt with until the university level. Space C contains six topics on relations and functions for which there are not corresponding test items, and space D four such calculus topics. Space G includes nine test items which do not deal directly with topics included in the matrix and guidelines, and are for the most part a review of the pre-Grade Thirteen work. Spaces A and B are the ones on which this analysis will concentrate, with some attention given to G. Space A includes ten topics from the matrix and eighteen related test items, while space B includes eight matrix topics and nine corresponding test items. It is in dealing with these spaces that it is possible to combine data from the two projects in the hope of deriving some information about the existence of program gaps and duplications.

One of the most important features of space C in particular is that almost all the topics included in it are included in the guidelines not only at the Grade Thirteen level but also at earlier levels. It is of some concern that these topics are not covered by the test. It is of more concern that space D contains topics which are first met at the Grade Thirteen level but are not tested. If a simple count of topics can be used as a rough guide, it appears that about 60% of the mandated (and widely taught) Grade Thirteen program in Relations and Functions and Calculus is covered by the Mathematics Achievement Test. It is also true that about one-quarter of the test deals with material not contained in those programs, but taught in earlier years. The amount and nature of the discrepancy between the program and the test is within reasonable limits, and the experience of the investigators suggests that it compares favourably with the match

between programs and current tests in other jurisdictions, both within and outside Canada.

2. ORGANIZATION OF THE DATA

The data on which the interproject analysis was based were first assembled in Topic Tables of the form illustrated in Appendix 4C. One of these tables was drawn up for each Project III topic for which there were corresponding Project II test items. Each table included the Project III data on the ratings given by teachers at both levels to student competence in the topic, and on the statements of these teachers about the level at which they would prefer student knowledge to be on entry to and on exit from their courses. (Only preliminary figures were available for Interproject Analysis; indications from the few final figures presented in the Project III report are that changes are small.) These teacher ratings were based on the following scale:

- 0 Has no knowledge of this topic
- 1 Has an awareness and descriptive knowledge of the topic only; has no ability to apply it.
- 2 Has elementary knowledge and skill, based on a simple quantitative introduction; can do simple examples and programs only.
- 3 Has a working grasp of the topic, and is able to do standard exercises and problems; has some facility in translating problems into mathematical terms, and in knowing the place of the topic in a wider context.
- 4 Has a thorough understanding of the topic and its place in mathematics, can solve a wider range of problems, prove theories, generalize and give examples.

- 5 Has complete mastery of the topic; underlying concepts are part of his mathematical background and he can apply them to solve non-routine, symbolic, verbal, or real-life problems. He understands the theoretical basis of the topic and its limitations of applicability.

For each related test item, the Topic Table contains the assessment by teachers at both levels, given on the Test Appraisal Inventory, of whether the item content should be known by entering students, should be taught during their course, or should not be known or dealt with at this level, along with a statement of the percentage of students currently answering the item. The assessments by secondary teachers were based on the following scale:

- A. Old knowledge that students should have on entry to the course.

A1. This knowledge is not reviewed in the course.

A2. This knowledge is reviewed in the course.

- B. New knowledge that all students are expected to learn in the course.

- C. New knowledge that some students are expected to learn in the course.

C1. Only 1% to 25% of students should learn this.

C2. Only 26% to 50% of students should learn this.

C3. Only 51% to 75% of students should learn this.

C4. More than 75% but not all students should learn this.

D. New knowledge that no student is expected to learn.

The scale used by university teachers was the same at the A and B levels, but the remaining secondary categories were replaced by:

C. Other.

The first important piece of information apparent from the Topic Tables was the consistent and large discrepancy between the secondary teachers' perception of the exit level of their students and the university teachers' perception of the level of their incoming students. The university teachers reported far lower student capability than did the secondary teachers. The pattern will be discussed as a general issue elsewhere in this chapter.

From the Topic Tables, a Summary Table (Table 4.1) was prepared. This table provides, for each topic for which there are related test items, the following information:

-- The mean reported secondary school exit level and university entrance level of competence in the topic.

-- For each related test item:

- (i) the assessment of the item content by teachers of Relations and Functions, teachers of Calculus, and first year university mathematics teachers. The way these assessments were arrived at is described below.
- (ii) the percentage of students correctly answering the item.
- (iii) a tentative judgement as to whether these data indicate the existence of a gap (shown by a G in the last column) or a duplication (shown by a D in the last column).

The indications in the Summary Table of teacher assessment of the content of each item were derived from the Test Appraisal Inventory data in the following way: For secondary teachers, the percentage of responses was accumulated in the order A1, A2, B, C4, C3, C2, C1, D, until 75% of the respondents were included. The point at which 75% was reached was then allocated to the item. This was done separately for Calculus teachers and Relations and Functions teachers. University teachers' responses were accumulated in the order A1, A2, B, C, until the total reached 70% of those responding. The slightly lower figure was used in this case to allow for the great variety of university courses and the consequently large number of university teachers assessing particular items as "C"--in most cases, not relevant to their particular course.

It should be pointed out that all data considered in this chapter were drawn from the part of the study dealing with Anglophone schools. The small numbers involved in the Francophone part of the study made it impossible to derive an accurate picture of the gap-overlap situation in Francophone institutions. There was, for example, only one French-language postsecondary respondent to the Test Appraisal Inventory.

Following a topic-by-topic discussion of the Summary Table, each item appearing to indicate the existence of a gap or a duplication is analyzed in more detail, with additional tabular data.

3. SUMMARY DATA

There is throughout the Summary Table a consistent and sizeable discrepancy between the perception of secondary teachers of the students' exit level from their classes on a given topic, and that of university teachers of the students' entry level on the same topic. This suggests that there has been a general lack of agreement between teachers at the two levels as to how the scale

was to be interpreted, and that little, if anything, can be deduced from the large discrepancy in a single case. What may be of some use, however, is an examination of the rank-ordering of these assessments, so that for a particular topic one can state that secondary teachers or university teachers considered competency in this topic high or low in comparison with competency in other topics. This is the practice that will be followed throughout this section; Section 6 presents some justification for the procedure.

3.1 Topic I-1: Function as a mapping

The rating by secondary teachers of competency in this topic is their eleventh highest; that by university teachers is their tenth highest. This correspondence is very close; and it is interesting that if we take as an index of student performance for each topic the mean of the percentages of students correctly answering the related items, the resulting 55 for Topic I-1 is the eleventh highest.

There are two items dealing with this topic. The Summary Table indicates that in each case both Calculus teachers and Relations and Functions teachers have assigned a B to the item suggesting coverage in both courses. The university assessment in each case is A2, previous knowledge reviewed at this level. The student success rate of 82% on item 7 indicates that there is no problem with the content of that item--it is taught and learned at the secondary level, and at most reviewed at the postsecondary level. However, the success rate on item 26 is very low--only 29%. This poor performance, combined with university expectations that this is previous knowledge, warrants a closer examination of the item in the next section of this chapter as possibly indicating a gap.

3.2 Topic I-2: Inverse of a function

Both university and secondary assessments of student competence on this topic are among the lowest three of their respective rank-orderings. This low rating is supported by student performance on the one related test item, which was correctly answered by only 30% of students. In spite of their low overall estimates of student competency with the topic, both sets of teachers indicated by their ratings of the item content that students should have covered the material before entering university. This item must tentatively be categorized as indicating a gap.

3.3 Topic I-3: Graphs and properties of second degree relations using previously known skills

There is a sharp difference for this topic between the rank-orderings of competence at the secondary and university levels. It is ranked third highest by university teachers, but last by secondary teachers. The explanation for this discrepancy may lie in the lack of specificity of the topic, or in the probably different interpretations at the two levels of the words "previously known skills".

Students performed well on one of the four related items, number 19. There seems to be no problem here, since performance was good and university teachers treat the item as previous knowledge. The same comment, with some reservations because of the somewhat lower success rate of 45%, applies to item 18. Item 20 clearly warrants further investigation as possibly indicating a gap, since only 36% of students answered the item correctly and most university teachers expected previous knowledge of its content. In light of this apparent problem, it seems wise to investigate further the responses to item 32, with a low success rate of 25%, to see whether the university rating of B is at all misleading.

3.4 Topic I-6: Applications (of equations of conics in non-standard positions)

Both university and secondary ratings of competence place this topic seventh highest. The related item has content widely covered in or before the Relations and Functions course, successfully learned (to judge by the 62% success rate), and yet assigned a B by university teachers. It may indicate the existence of a duplication.

3.5 Topic I-8: Intersection of conics and conics

Secondary teachers and university teachers disagree sharply in their relative rankings of student competence here. University teachers place it near the bottom of their list, in fourteenth place, while secondary teachers rank it fifth. The two relevant test items have student success rates which are also somewhat contradictory--a high rate of 80% on one item, and a low one of 34% on the other. In the case of this second item, the university assessment of its content as previous knowledge suggests the existence of a gap.

3.6 Topic I-10: Standard trigonometric formulae and applications

The secondary school competency rating for this topic is near the bottom of the ordering; the university rating is fourth highest. The relatively low success rates on the four related items lend support to the secondary rating as closer to reality. There is certainly a perceptual gap here, and there may well be a real one in view of university expectations of previous knowledge of the content of at least three of these four items. The problem is clearest with item 28, with a 28% success rate and an A2 rating by university teachers. But the other three items are close enough to the borderline of the problem area to merit further study.

3.7 Topic I-12: Phase shift, period and amplitude

Both sets of teachers put student competency in this topic high on their lists--fourth in the case of secondary teachers, and eighth for university teachers. The one related item was very well done, with a success rate of 80%, and its content was at most reviewed at university. There was no indication of either gap or duplication here.

3.8 Topic I-13: Translations of the plane

There is a sharp difference of opinion here between secondary and university teachers about relative student competency, which is not resolved by examining the performance data. Secondary teachers put this topic third highest; university teachers ranked it thirteenth; and the student performance was seventh best of that on any topic. Although it is not suggested by the summary of data dealing with the particular item, there is certainly a possibility of duplication here, given the high secondary estimate of competency and the good level of performance (62% correct) on the one hand, and the low university estimate on the other.

3.9 Topic I-14: Rotations of the plane

Competency in this topic was ranked ninth and fifteenth by secondary and university teachers respectively; student performance ranked twelfth among topics. In spite of somewhat differing perceptions at the two levels, this does not appear to be a problem area; student performance was good (52% correct), and the C rating by university teachers indicates that there are many courses in which the topic is not included.

3.10 Topic I-15: Reflections of the plane

University and secondary teachers respectively rated competency on this topic eighth and sixteenth--a marked difference. Student performance was good, with a 67% success rate, but in any case the C rating of the item content suggests that the gap or duplication problem does not arise with this item.

3.11 Topic II-2: Rate of change: slopes, secants, tangents

This topic comes fifth in the competency rankings of both sets of teachers. Mean student performance on the two related test items ranks eighth among topics. In spite of good performance on both these items (77% and 49%), the university ratings of the content of both is B, raising the question of duplication for investigation.

3.12 Topic II-3: Derivatives of powers, products and quotients

This topic is ranked first in terms of competency by both secondary and university teachers, and student performance on the one related item is the highest on any test item, with 92% success. In spite of all these facts, university teachers appear to reteach the item content. This is a very probable case of duplication.

3.13 Topic II-5: Applications of derivatives to tangents to curves

Teachers at the secondary level appeared to be reasonably well satisfied with student performance on this topic, ranking it sixth. University teachers put it only slightly lower, in ninth place. However, the test results contradict these respectable rankings, with only 25% of students correctly answering the one item dealing with the topic--one of the three lowest results for

any topic. The B rating by university teachers appears to result from a justified reteaching of unlearned content, rather than to be an instance of duplication.

3.14 Topic II-6: Further applications (of derivatives): velocity, acceleration

The rank-ordering of competency ratings places this topic ninth on the secondary teachers' list and eleventh on that of the university teachers. Actual test performance ranks somewhat lower, at fourteenth; the one related item was correctly answered by only 45% of students. The university assessment of the test content as C suggests that this topic is not included in many first year courses. In any case, the 45% success rate is too high to indicate a gap, and too low to suggest a duplication of effort, whatever the university treatment of the topic might be.

3.15 Topic II-8: Maxima and minima problems

This topic ranked low relative to the others on both competency scales; it was listed as fifteenth by secondary teachers, and as twelfth by university teachers. Student performance on the related item was very good, at 70% correct. The B assigned by university teachers to the content of this item, in combination with the high success rate, warrants a tentative categorization as duplication.

3.16 Topic II-9: Rate of change problems

The competency rankings for this topic are low--seventeenth for secondary teachers, fourteenth for university teachers--and the poor success rate of 23% on the one item dealing with the topic supports those perceptions. The indication of reteaching at the university level is thus not interpretable as suggesting a

duplication; it appears to be a necessary second coverage of material not well mastered in Grade Thirteen.

3.17 Topic II-12: Areas enclosed between curves

Competency in this topic was also ranked low by both sets of teachers--thirteenth by secondary teachers, eighteenth (or last) by university teachers. It is not surprising then that only 7% of the students correctly answered the one test item on the topic, and that the material appears from the Summary Table to be retaught at the university level.

One peculiarity of this test item should perhaps be pointed out. It may be (and, to judge by the test results, probably was) attacked as a standard problem involving definite integrals. However, it is also susceptible to a second method of solution involving only very rudimentary Euclidean geometry. It is clear that few, if any, students (or teachers) considered the possibility of this alternate approach, which made the item considerably less difficult and complex than did the standard method.

3.18 Topic II-28: Derivatives of elementary functions

This is the only topic from the second part of the "Calculus" section of the content matrix which is represented by a test item. This second part consists of topics treated at an advanced rather than an elementary level. Competency in this topic was placed second on the list by both groups of teachers, and student success on the test item was a very high 89%. However, the B rating assigned the item content at the university level raises the possibility of some duplication here.

3.19 Test items not matched to topics

Although the central purpose of this analysis is to compare Project II and Project III data for the purpose of identifying program gaps and duplications between the Grade Thirteen and university levels, it may be of some interest to look at the nine test items not clearly identifiable with any of the topics included in the content matrix. There is, of course, only Project II data available for these items. All these items were considered carefully by the project teams, and evaluated as reviewing pre-Grade Thirteen work, largely in the area of relations and functions, although there is one item dealing with each of algebra, synthetic geometry and trigonometry.

The data in the Summary Table reveal three items--numbers 11, 14 and 31--where a gap may exist. In all three cases, the student success rate was 32% or lower. In two of these cases, the universities clearly expected the students to have learned the item content. The university assessment of content in the third case is a C, which may mean that this material is not important at that level, but it seems worth examining this item along with the other two. Two duplications are also suggested by items 22 and 29. These items were answered correctly by 53% and 51% of students respectively, but received a B rating from the university teachers.

4. GAPS

In the discussion of gaps in this section, and that of duplications in the next section, some additional data will be examined along with that gathered on a provincial scale. It is possible that, even though the existence of a gap or duplication over the province as a whole is confirmed, on a local scale the situation might be different. Consequently, two somewhat limited case studies were carried out in connection with the items tentatively assigned to the category of gap or duplication.

In the first instance, a group of three institutions was considered--one university, where there were five respondents to the Test Appraisal Inventory, and two nearby secondary schools, with two and three respondents respectively. In the first of these schools, one respondent taught Calculus and the other Relations and Functions. In the second school, there were two Relations and Functions teachers responding, and one Calculus teacher.

The second case study dealt with one university, providing five respondents, and one nearby secondary school with three respondents, two teaching Relations and Functions and one teaching Calculus.

These case study data are admittedly limited, and can serve at best only to support or cast doubt on the hypothesis that province-wide gaps and duplications are reflected at the local level. As well as the problem of the small sample size, which severely affects the reliability of the data, there is a second problem in that difficulty indices for the students in the case study schools are not available. So what evidence there is bears only on teaching patterns, and not on student performance.

The analysis done in the preceding section indicates that there are few if any gaps which arise from lack of coverage of a topic at the secondary level. What may exist, and will be the main object of search in this section, are gaps such that students, although they have covered the material, display a lack of proficiency in topics which their university instructors assume they know adequately.

4.1 Topic I-1: Function as a mapping

Table 4.2 presents detailed data on Test Appraisal Inventory responses to the relevant item, both on a province-wide scale and for the case study institutions individually. 72% of Relations and Functions teachers and the same percentage of Calculus

teachers considered the content of this item to be already old knowledge at the Grade Thirteen level. 80% of university teachers considered it previous knowledge, and 31% did not even review the material. In this context, the 29% success rate is startling. The type of gap indicated here might be called a "super-gap", since it apparently has its origins before the final secondary year, and is not filled during that year.

The pattern is similar in both case studies. The five mathematics teachers in the Case Study 1 schools, with one exception, assumed this content to have been learned before Grade Thirteen, and all four university teachers who deal with the topic treated it as old knowledge. The three Case Study 2 secondary teachers all handled it as review material, and three of the four university teachers did likewise.

There is clearly a super-gap in this topic. However, this is not true of all parts of the topic, as is witnessed by the excellent student performance on item 7.

4.2 Topic I-2: Inverse of a function

Data on test item 8, the only one on this topic, are presented in Table 4.3. This item appears to be a second example of the "super-gap" problem. At the Grade Thirteen level, 52% of Relations and Functions teachers and 68% of Calculus teachers treated the item content as previous knowledge. A sizeable percentage of the former did handle it as new material, so the super-gap status is not so clear as with the previous topic. 71% of university teachers expected students to have learned this material before entering their courses.

In Case Study 1, the material was considered old by three secondary teachers and taught as new by the remaining two. All three university teachers dealing with the material did not even review it. The Case Study 2 secondary teachers unanimously

regarded it as previous knowledge, as did three of the four university teachers.

The question with this topic is not whether there is a gap--that seems evident--but rather at what point the gap arises. At least in part, responsibility seems to lie at the lower levels of the secondary schools.

4.3 Topic I-3: Graphs and properties of second degree relations using previously known skills

Table 4.4 contains data on the two test items of four on this topic which raise the question of the existence of a gap. The response pattern for item 20 is somewhat peculiar. The content appears to belong primarily to the Relations and Functions course, since it is taught or assumed by 93% of teachers of that course. Yet only 3% of these teachers considered it old knowledge, while 30% of the Calculus teachers put it in this category. At any rate, it had been covered by the above-mentioned 93% of Calculus teachers by the end of the Grade Thirteen year, and was assumed to be old knowledge by virtually all university teachers dealing with the topic.

All three Relations and Functions teachers in Case Study 1 taught this content, and all three university teachers dealing with the topic assumed previous knowledge. Not only both Relations and Functions teachers, but the Calculus teacher as well in Case Study 2 taught the material, and it was assumed to be known by all three interested university teachers.

The data on item 32 indicate that there is not in fact a gap connected with this item. The low degree of coverage in secondary school, and the consequent poor student performance, are not important in view of the small proportion of university teachers who expect this knowledge on entry. The case study data are somewhat different from the provincial data in this case, in suggesting both a higher degree of coverage in the secondary

courses and a higher level of expectation at the university level.

One of four items on this topic, then, points out a gap, which for the most part originates at the Grade Thirteen level. The variations in performance and response patterns among these items are probably a result of the breadth of the topic and a consequent lack of homogeneity among the items.

4.4 Topic I-8: Intersection of conics and conics

Table 4.5 presents the available data on test item 30. The content was widely taught in the Relations and Functions course, and expected as previous knowledge by 71% of university teachers. This expectation was not met by the 34% success rate on the item.

The Case Study 2 data fit the provincial pattern, with all three secondary teachers covering the material. Two of the three interested university teachers expected previous knowledge.

In Case Study 1, there is sharp variation. Only one of the three Relations and Functions teachers taught this content to all (or even half) the students, though again two of the three university teachers dealing with the material treated it as old knowledge.

A gap is confirmed on this item content, but it is not topic-wide, as may be seen by student performance on item 1.

4.5 Topic I-10: Standard trigonometric formulae and applications

Only one of the four items on this topic clearly indicates the possibility of a gap. However, the other three show a borderline pattern in the Summary Table, and it seems wise to examine all four; detailed data are given in Table 4.6.

The content of each of these four items was expected at university entrance by at least 69% of teachers at that level. In all four cases, the content was considered previous knowledge at the Grade Thirteen level by sizeable numbers of secondary teachers. Student performance on one was poor, and on the other three just adequate. These data confirm the diagnosis of a gap in this topic, though not as serious a one as some already discussed in that the average performance level was close to 40%. It may be more serious in another way, since it is manifested in all four items. The evidence is strong, particularly in the cases of items 4 and 23, that what exists here is another super-gap leading back beyond Grade Thirteen to inadequate learning at earlier levels.

The case study data fit the general pattern for items 4 and 28, with some variation in item 27 for Case Study 2, and in item 23 for both case studies. Here the high university expectations were manifested, but there was much less secondary coverage of the topic.

4.6 Pre-Grade Thirteen Items Suggesting Gaps

Table 4.7 details inventory responses to the three test items on the pre-Grade Thirteen level where gaps are indicated. The content of these items is varied; they review respectively topics in functions, synthetic geometry and analytic geometry. In every case, secondary teachers agreed overwhelmingly with their categorization as earlier work. Even allowing for a number of university teachers who did not deal with the material (from 13% to 38%), university expectations of student knowledge were very high; from 49% to 62% of university teachers did not even review the material. There seems no doubt that three super-gaps exist in the areas covered by these items. Both local pictures are very similar to the provincial one, with almost unanimous expectations at both Grade Thirteen and university levels that students would know this material on entry.

4.7 Summary

The suspicions aroused by the twelve items discussed above have proven to be justified in all but one case. The reliability of the evidence is somewhat variable, ranging from one extreme at which only one test item of four on a topic shows the problem, to the other at which it is evidenced in all four of the items dealing with a second topic. The most dubious cases are those in which a topic has only one related item, and undue weight may be put on the test results on that item.

What is consistent and provocative about these findings, however, is that almost without exception they occur in the domain of Relations and Functions, and that in nine of the eleven cases, the gap revealed by an item has had its origins before the Grade Thirteen level--what we have called a super-gap. It appears that secondary teachers at this level place an undue amount of faith in the knowledge of their incoming students, and consequently perpetuate their lack of knowledge by not reteaching material on the assumption that students have already mastered it.

5. DUPLICATIONS

There are nine test items for which the data in the Summary Table suggest that their content is unnecessarily retaught at the university level. As in the previous section, each of these items will be studied more closely in order to confirm or deny that tentative judgement. Each of these items was successfully answered by 50% or more of the students, and each has a summary rating by university teachers of B (or, in one case, C).

5.1 Topic I-6: Applications (of equations of conics in non-standard positions)

Table 4.8 gives detailed provincial and case study data on item 12, the only one on this topic. The problem of duplication is seen, in the light of this table, to be real; 15% of university teachers treated the material as completely new, and 38% spent probably unnecessary time reviewing it.

The picture is similar in both case studies, with extensive secondary coverage or review, and for the most part review at the university level.

5.2 Topic I-13: Translations of the plane

Table 4.9, which contains data on test item 15, suggests that the duplication in this topic is illusory. 15% of university teachers handled the item content as new knowledge, but the percentage reviewing it was smaller than in the previous case. The Case Study 1 data fit this general picture, although one Case Study 2 university teacher retaught the material, which had been less extensively taught in the Relations and Functions course in this case.

5.3 Topic II-2: Rate of change: slopes, secants, tangents

Table 4.10 gives details on the inventory responses to both test items on this topic, both of which are suggestive of duplication. In both cases, there is an increase over the items previously discussed in the percentage of university teachers presenting the material as new, and an even sharper increase in the number reviewing it. In the case of item 9, we find that 69% of these teachers were reviewing or teaching material in which students were competent enough to score a 77% success rate on entry. In the two case studies, this time was spent only on review, but even this seems unnecessary duplication of effort.

5.4 Topic II-3: Derivatives of powers, products and quotients

Item 13, the only item on this topic, is covered in Table 4.11. This may be the most striking instance of duplication encountered in this analysis. 67% of university teachers reviewed or taught material on which the success rate was 92% (although only 23% treat it as entirely new). The content was covered by every Calculus teacher involved in the study. The situation is somewhat less pronounced in the case study universities, where no teacher treated the material as new, but this implies that there was a greater amount of duplication in other universities.

5.5 Topic II-8: Maxima and minima problems

Detailed data on item 24, the sole representative of this topic, appear in Table 4.12. Of university teachers dealing with the topic, none assumed this knowledge without review, and half taught it as new. In view of the 70% success rate on the item, this seems like largely wasted effort. The general picture holds in the case study institutions as well.

5.6 Topic II-28: Derivatives of elementary functions

Table 4.13, which deals with item 10, reveals that the comments about the previous topic apply largely here as well. Student performance was excellent, yet no university teacher treating the topic failed to review or reteach the material. The same, of course, was true in the target universities.

5.7 Pre-Grade Thirteen Items Suggesting Duplications

Table 4.14 includes the Project II data on test items 22 and 29, which review earlier work on functions. The duplication problem is not present with item 29; the great majority of university teachers did not teach or review the item content, but

assumed knowledge of it. Over half did, however, teach or review the content of item 22, although in most cases what took place was review. There appears to be a minor duplication of effort here. The general pattern for both these items holds at the case study level.

5.8 Summary

Nine items have been looked at for confirmation or denial of the hypothesis that they indicate program duplication. In seven of these cases, the hypothesis has been supported.

In the same way that almost all gaps occurred in the topics falling into the domain of relations and functions, all but two of these duplications occurred in the area of calculus.

6. CONFORMITY OF PROJECT II DATA WITH PROJECT III DATA

One of the difficulties in carrying out interproject analysis has been the substantial discrepancy between the rating standards of secondary and university teachers when asked to assess the competence of students on a particular topic. The secondary ratings are consistently much higher than the university ones. In an effort to derive some meaning from this apparently discrepant data, the topics were rank-ordered first by secondary teachers' assessments and then by those of university teachers. What emerged from that ranking is summarized in Table 4.15. It is clear that there is positive correlation between student performance and each of these sets of ratings. It would appear, then, that the discrepancies between the ratings are consistent across topics, and that they are related to the empirical performance data.

It is of further significance that the same two topics are ranked highest on both lists, and that both displayed the problem of duplication. Further, the one item occurring among the bottom three on both lists proved to be an instance of a gap. Such confirmations as these cannot be taken as rigid indicators of reliability or consistency across the data from the two projects, but they do suggest that some degree of such reliability and consistency is present.

7. SUMMARY

The Project II report (Traub et al., 1976) contains as Appendix A6 a technical report on the administration of the Mathematics Achievement Test, and comparative data on its earlier use (1968) as the Ontario Mathematics Achievement Test, Form OB. Although, for reasons outlined in that report, some caution must be used in making historical comparisons because of curriculum changes and certain differences in the populations, the evidence is strong that student performance on the test has been relatively stable over the two administrations. What this seems to indicate is that present curriculum problems in mathematics can not be solved by a return to earlier secondary programs. A number of gaps and duplications have been identified in this analysis, and although the list is certainly not exhaustive because the test does not provide complete coverage of the program, these data form a basis for further investigation and decision-making.

The test employed in this study is, of course, not ideal for measuring student achievement in Ontario Grade Thirteen mathematics courses. One-quarter of the test items deal with material covered in the curriculum only before the Grade Thirteen level, and it could be argued that treatment of such material is inappropriate in a test at this level. Coverage is also given in the field of relations and functions to material treated both before and during Grade Thirteen. However, particularly in view of the number of cases in which such items revealed inadequate

student knowledge in the present study, it does in fact seem important to test this sort of material.

A more serious problem with this test is its lack of coverage of a number of topics dealt with in Calculus, and of the entire content of Algebra. These areas should certainly be dealt with in any future standardized testing program. It would also be wise, in view of differential enrolment in these courses, to use separate tests for each course. In this study, for example, it was not possible to test students enrolled in only one of Calculus and Relations and Functions, since the test material could not be separated.

An interesting problem of interpretation is raised by the fact that Grade Thirteen teachers involved in this study consistently estimated the level of performance of their students on exit in a particular topic as much higher than what was indicated by university teachers to be the entrance level of those students to university courses. Logically, one would expect any discrepancy to work in the opposite direction, since those students entering university from Grade Thirteen may be assumed to be, on the average, somewhat more capable than those not proceeding to further studies.

One way of dealing with this anomaly in the data is to suggest that these ratings are unreliable, and should be excluded from further analyses. However, if the university ratings are converted to ranks, and the same procedure is carried out on the secondary ratings, there is a close correspondence of these rankings not only with one another but also with the actual performance of students on test items dealing with the topics in question. (It is interesting, though not surprising, that the correspondence with performance is closer in the case of secondary teachers than it is for university teachers.) These data are of use, then, in the altered form of rank-orderings.

A number of gaps have been pointed out by items on which the student success rate was less than 40% (or slightly above this in certain special cases), and on which a sizeable percentage of university teachers expected incoming students to have a knowledge of the content. These gaps are evidenced in the following topics:

- (a) I-1: Function as a mapping (1 item of 2)
- (b) I-2: Inverse of a function (1 item of 1)
- (c) I-3: Graphs and properties of second degree relations using previously known skills (1 item of 4)
- (d) I-8: Intersection of conics and conics (1 item of 2)
- (e) I-10: Standard trigonometric formulae and applications (4 items of 4)
- (f) Pre-Grade Thirteen material (3 items of 9)

Duplications have been indicated by items where the success rate was 49% or better, and where significant numbers of university teachers stated that they reteach the topic or perform apparently unnecessary review on it. Duplications occurred in the following topics:

- (a) I:6 Applications (of equations of conics in non-standard positions) (1 item of 1)
- (b) II-2: Rate of change: slopes, secants, tangents (2 items of 2)
- (c) II-3: Derivatives of powers, products, and quotients (1 item of 1)

(d) II-8: Maxima and minima problems (1 item of 1)

(e) II-28: Derivatives of elementary functions
(1 item of 1)

(f) Pre-Grade Thirteen material (1 item of 9)

It is important that most of the duplications involve calculus items, and most of the gaps are related to topics in relations and functions. It is perhaps even more important that almost all gaps could be traced back to levels before Grade Thirteen.

Some approaches at the secondary level to achieving more efficient learning in mathematics suggested by this analysis, then, are a greater emphasis at pre-Grade Thirteen levels on mastery of those skills needed in the senior courses; a more careful examination by Grade Thirteen teachers of the extent of students' previous knowledge; and more effort by these teachers to remedy any existing gaps in basic knowledge. At the university level, teachers might safely put more confidence in their students' knowledge of calculus, and use the time saved in this way in more extensive review and teaching of those areas in which student competence appears to be lower.

TABLE 4.1

SUMMARY OF TOPIC TABLES

Topic	Reported secondary exit level*	Reported university entrance level*	Assessment by**				% correct	Suggested gap or duplication
			Item No.	R and F teachers	Calculus teachers	university teachers		
I-1	3.20	1.65	7	B	B	A2	82	-
			26	B	B	A2	29	G
I-2	3.06	1.11	8	B	B	A2	30	G
I-3	2.90	1.91	18	B	A2	A2	45	-
			19	A2	A2	A2	78	-
			20	B	D	A2	36	G
			32	D	C2	B	25	G
I-6	3.31	1.80	12	B	C4	B	62	D
I-8	3.13	1.81	1	B	B	A1	80	-
			30	B	D	A2	34	G
I-10	3.20	1.83	4	B	B	A2	42	-
			23	D	B	B	45	-
			27	B	B	A2	42	-
			28	B	B	A2	28	G
I-12	3.42	1.75	6	A2	B	A2	80	-
I-13	3.49	1.50	15	B	D	C	62	D

TABLE 4.1 (continued)

Topic	Reported secondary exit level*	Reported university entrance level*	Item No.	Assessment by**			% correct	Suggested gap or duplication
				R and F teachers	Calculus teachers	university teachers		
I-14	3.22	1.20	16	B	D	C	52	-
I-15	3.30	1.20	17	C3	D	C	67	-
II-2	3.39	1.81	9	D	B	B	77	D
			33	D	B	B	49	D
II-3	3.69	2.04	13	D	B	B	92	D
II-5	3.37	1.73	25	C4	B	B	25	-
II-6	3.22	1.60	34	D	B	C	45	-
II-8	3.08	1.55	24	D	B	B	70	D
II-9	3.04	1.47	35	D	B	B	23	-
II-12	3.15	0.83	36	D	C	B	7	-
II-28	3.67	2.00	10	D	B	B	89	D

TABLE 4.1 (continued)

Topic	Reported secondary exit level*	Reported university entrance level*	Assessment by**				% correct	Suggested gap or duplication
			Item No.	R and F		university teachers		
				teachers	Calculus teachers			
Pre-Grade Thirteen work			2	B	D	C	55	-
			3	A2	A2	A2	85	-
			5	A2	A2	A1	57	-
			11	A2	A2	A2	32	G
			14	A2	A2	C	12	G
			21	A2	B	A2	67	-
			22	A2	A2	B	53	D
			29	A2	A2	B	51	D
			31	A2	A2	A2	18	G

*Preliminary figures. See text, Section 2, for explanation and definition of scale.

**See text, Section 2, for explanation of assessments.

TABLE 4.2

DETAILED DATA ON TOPIC I-1: FUNCTION AS A MAPPING
AND RELATED TEST ITEM 26

(a) Province-wide data--Inventory responses* and student performance

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers (%)	6	66	23	0	1	1	2	1
<u>Calculus</u> teachers (%)	34	38	14	0	0	0	2	12
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University teachers (%)	31	49	15	5	% correct = 29			

(b) Case study data--Inventory responses*

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers, school 1-1	0	1	0	0	0	0	0	0
<u>Calculus</u> teachers, school 1-1	0	1	0	0	0	0	0	0
<u>R and F</u> teachers, school 1-2	0	1	0	0	1	0	0	0
<u>Calculus</u> teachers, school 1-2	1	0	0	0	0	0	0	0
<u>R and F</u> teachers, school 2	0	2	0	0	0	0	0	0
<u>Calculus</u> teachers, school 2	0	1	0	0	0	0	0	0
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University 1	3	1	0	1				
University 2	1	2	1	0				

*See text, Section 2.

TABLE 4.3

DETAILED DATA ON TOPIC I-2: INVERSE OF A FUNCTION
AND RELATED TEST ITEM 8

(a) Province-wide data--Inventory responses* and student performance

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers (%)	2	50	43	1	2	0	1	1
<u>Calculus</u> teachers (%)	34	34	15	0	0	0	2	15
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University teachers (%)	33	38	18	10	%correct = 30			

(b) Case study data--Inventory responses*

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers, school 1-1	0	0	1	0	0	0	0	0
<u>Calculus</u> teachers, school 1-1	1	0	0	0	0	0	0	0
<u>R and F</u> teachers, school 1-2	0	1	1	0	0	0	0	0
<u>Calculus</u> teachers, school 1-2	1	0	0	0	0	0	0	0
<u>R and F</u> teachers, school 2	0	2	0	0	0	0	0	0
<u>Calculus</u> teachers, school 2	0	1	0	0	0	0	0	0
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University 1	3	0	0	2				
University 2	1	2	1	0				

*See text, Section 2.

TABLE 4.4

DETAILED DATA ON TOPIC I-3; GRAPHS AND PROPERTIES OF SECOND
DEGREE RELATIONS USING PREVIOUSLY KNOWN SKILLS
 AND RELATED TEST ITEMS 20 AND 32

ITEM 20 FIGURES

(a) Province-wide data--Inventory responses* and student
 performance

	A1	A2	B	C4	C3	C2	C1	D
<u>R and F</u> teachers (%)	2	1	90	2	1	2	2	0
<u>Calculus</u> teachers (%)	16	23	19	1	0	3	1	36
	A1	A2	B	C				
University teachers (%)	59	18	3	20	% correct = 36			

(b) Case study data--Inventory responses*

	A1	A2	B	C4	C3	C2	C1	D
<u>R and F</u> teachers, school 1-1	0	0	1	0	0	0	0	0
<u>Calculus</u> teachers, school 1-1	0	0	0	0	0	0	0	1
<u>R and F</u> teachers, school 1-2	0	0	2	0	0	0	0	0
<u>Calculus</u> teachers, school 1-2	1	0	0	0	0	0	0	0
<u>R and F</u> teachers, school 2	0	0	2	0	0	0	0	0
<u>Calculus</u> teachers, school 2	0	0	1	0	0	0	0	0
	A1	A2	B	C				
University 1	2	1	0	2				
University 2	2	1	0	1				

TABLE 4.4 (continued)

ITEM 32 FIGURES

(a) Province-wide data--Inventory responses* and student performance

	A1	A2	B	C4	C3	C2	C1	D
<u>R and F</u> teachers (%)	13	13	24	0	1	3	11	36
<u>Calculus</u> teachers (%)	9	4	48	6	4	6	10	12
	A1	A2	B	C				
University teachers (%)	8	28	38	26	% correct = 25			

(b) Case study data--Inventory responses*

	A1	A2	B	C4	C3	C2	C1	D
<u>R and F</u> teachers, school 1-1	1	0	0	0	0	0	0	0
<u>Calculus</u> teachers, school 1-1	0	0	1	0	0	0	0	0
<u>R and F</u> teachers, school 1-2	0	1	0	0	0	0	0	1
<u>Calculus</u> teachers, school 1-2	0	0	1	0	0	0	0	0
<u>R and F</u> teachers, school 2	0	0	1	0	0	1	0	0
<u>Calculus</u> teachers, school 2	0	0	1	0	0	0	0	0
	A1	A2	B	C				
University 1	1	1	0	3				
University 2	1	2	0	1				

*See text, Section 2.

TABLE 4.5

DETAILED DATA ON TOPIC I-8: INTERSECTION OF CONICS AND CONICS
AND RELATED TEST ITEM 30

(a) Province-wide data--Inventory responses* and student performance

	A1	A2	B	C4	C3	C2	C1	D
<u>R and F</u> teachers (%)	1	2	76	3	3	9	5	1
<u>Calculus</u> teachers (%)	10	16	18	2	0	3	5	46
	A1	A2	B	C				
University teachers (%)	51	20	5	23	% correct = 34			

(b) Case study data--Inventory responses*

	A1	A2	B	C4	C3	C2	C1	D
<u>R and F</u> teachers, school 1-1	0	0	1	0	0	0	0	0
<u>Calculus</u> teachers, school 1-1	0	0	0	0	0	0	0	1
<u>R and F</u> teachers, school 1-2	0	0	0	0	0	1	1	0
<u>Calculus</u> teachers, school 1-2	1	0	0	0	0	0	0	0
<u>R and F</u> teachers, school 2	0	0	2	0	0	0	0	0
<u>Calculus</u> teachers, school 2	0	0	1	0	0	0	0	0
	A1	A2	B	C				
University 1	2	0	1	2				
University 2	1	1	1	1				

*See text, Section 2.

TABLE 4.6

DETAILED DATA ON TOPIC I-10:
STANDARD TRIGONOMETRIC FORMULAE AND APPLICATIONS
 AND RELATED TEST ITEMS 4, 23, 27, 28

ITEM 4 FIGURES

(a) Province-wide data--Inventory responses* and student performance

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers (%)	8	61	20	2	3	1	0	5
<u>Calculus</u> teachers (%)	23	45	11	1	1	1	0	17
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University teachers (%)	54	26	3	18	% correct = 42			

(b) Case study data--Inventory responses*

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers, school 1-1	0	1	0	0	0	0	0	0
<u>Calculus</u> teachers, school 1-1	0	0	0	1	0	0	0	0
<u>R and F</u> teachers, school 1-2	0	1	1	0	0	0	0	0
<u>Calculus</u> teachers, school 1-2	0	1	0	0	0	0	0	0
<u>R and F</u> teachers, school 2	0	1	0	1	0	0	0	0
<u>Calculus</u> teachers, school 2	0	1	0	0	0	0	0	0
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University 1	3	0	0	2				
University 2	2	1	0	1				

TABLE 4.6 (continued)

ITEM 23 FIGURES

(a) Province-wide data--Inventory responses* and student performance

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers (%)	34	21	10	0	2	1	3	30
<u>Calculus</u> teachers (%)	21	44	14	0	2	0	0	20
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University teachers (%)	20	49	5	26	% correct = 45			

(b) Case study data--Inventory responses*

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers, school 1-1	1	0	0	0	0	0	0	0
<u>Calculus</u> teachers, school 1-1	0	1	0	0	0	0	0	0
<u>R and F</u> teachers, school 1-2	0	0	0	0	0	1	1	0
<u>Calculus</u> teachers, school 1-2	0	1	0	0	0	0	0	0
<u>R and F</u> teachers, school 2	0	0	0	0	0	0	1	1
<u>Calculus</u> teachers, school 2	0	0	1	0	0	0	0	0
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University 1	0	2	0	3				
University 2	1	2	0	1				

TABLE 4.6 (continued)

ITEM 27 FIGURES

(a) Province-wide data--Inventory responses* and student performance

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers (%)	2	36	53	3	2	1	1	2
<u>Calculus</u> teachers (%)	24	46	16	0	0	0	0	14
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University teachers (%)	26	54	0	20	% correct = 42			

(b) Case study data--Inventory responses*

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers, school 1-1	0	1	0	0	0	0	0	0
<u>Calculus</u> teachers, school 1-1	1	0	0	0	0	0	0	0
<u>R and F</u> teachers, school 1-2	0	0	0	0	1	1	0	0
<u>Calculus</u> teachers, school 1-2	1	0	0	0	0	0	0	0
<u>R and F</u> teachers, school 2	0	1	0	1	0	0	0	0
<u>Calculus</u> teachers, school 2	0	0	1	0	0	0	0	0
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University 1	0	3	0	2				
University 2	1	2	1	0				

TABLE 4.6 (continued)

ITEM 28 FIGURES

(a) Province-wide data--Inventory responses* and student performance

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers (%)	4	43	44	2	2	1	0	3
<u>Calculus</u> teachers (%)	33	28	16	2	0	2	2	19
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University teachers (%)	38	41	3	18	% correct = 28			

(b) Case study data--Inventory responses*

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers, school 1-1	0	1	0	0	0	0	0	0
<u>Calculus</u> teachers, school 1-1	1	0	0	0	0	0	0	0
<u>R and F</u> teachers, school 1-2	0	1	1	0	0	0	0	0
<u>Calculus</u> teachers, school 1-2	1	0	0	0	0	0	0	0
<u>R and F</u> teachers, school 2	0	1	0	1	0	0	0	0
<u>Calculus</u> teachers, school 2	0	1	0	0	0	0	0	0
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University 1	1	2	0	2				
University 2	2	1	0	1				

*See text, Section 2.

TABLE 4.7

DETAILED DATA ON PRE-GRADE THIRTEEN ITEMS 11, 14 AND 31

ITEM 11 FIGURES

(a) Province-wide data--Inventory responses* and student performance

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers (%)	35	41	6	4	1	3	6	3
<u>Calculus</u> teachers (%)	45	32	11	0	0	5	2	6
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University teachers (%)	62	15	3	20	% correct =32			

(b) Case study data--Inventory responses*

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers, school 1-1	1	0	0	0	0	0	0	0
<u>Calculus</u> teachers, school 1-1	1	0	0	0	0	0	0	0
<u>R and F</u> teachers, school 1-2	1	0	0	0	0	0	1	0
<u>Calculus</u> teachers, school 1-2	1	0	0	0	0	0	0	0
<u>R and F</u> teachers, school 2	0	1	0	1	0	0	0	0
<u>Calculus</u> teachers, school 2	0	1	0	0	0	0	0	0
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University 1	3	0	0	2				
University 2	2	1	0	1				

TABLE 4.7 (continued)

ITEM 31 FIGURES

(a) Province-wide data--Inventory responses* and student performance

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers (%)	40	45	10	0	1	1	0	2
<u>Calculus</u> teachers (%)	46	37	9	0	0	0	2	6
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University teachers (%)	54	26	8	13	% correct = 18			

(b) Case study data--Inventory responses*

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers, school 1-1	1	0	0	0	0	0	0	0
<u>Calculus</u> teachers, school 1-1	1	0	0	0	0	0	0	0
<u>R and F</u> teachers, school 1-2	0	2	0	0	0	0	0	0
<u>Calculus</u> teachers, school 1-2	1	0	0	0	0	0	0	0
<u>R and F</u> teachers, school 2	0	2	0	0	0	0	0	0
<u>Calculus</u> teachers, school 2	0	0	1	0	0	0	0	0
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University 1	1	2	0	2				
University 2	3	1	0	0				

*See text, Section 2.

TABLE 4.8

DETAILED DATA ON TOPIC I-6: APPLICATIONS (OF EQUATIONS
OF CONICS IN NON-STANDARD POSITIONS)
AND RELATED TEST ITEM 12

(a) Province-wide data--Inventory responses* and student performance

	A1	A2	B	C4	C3	C2	C1	D
<u>R and F</u> teachers (%)	16	30	33	2	3	6	2	8
<u>Calculus</u> teachers (%)	26	24	23	1	1	1	9	14
	A1	A2	B	C				
University teachers (%)	20	38	15	26	%correct =62			

(b) Case study data--Inventory responses*

	A1	A2	B	C4	C3	C2	C1	D
<u>R and F</u> teachers, school 1-1	0	1	0	0	0	0	0	0
<u>Calculus</u> teachers, school 1-1	0	1	0	0	0	0	0	0
<u>R and F</u> teachers, school 1-2	0	0	1	0	1	0	0	0
<u>Calculus</u> teachers, school 1-2	1	0	0	0	0	0	0	0
<u>R and F</u> teachers, school 2	0	1	1	0	0	0	0	0
<u>Calculus</u> teachers, school 2	0	0	1	0	0	0	0	0
	A1	A2	B	C				
University 1	1	2	0	2				
University 2	1	2	0	1				

*See text, Section 2.

TABLE 4.9

DETAILED DATA ON TOPIC I-13: TRANSLATIONS OF THE PLANE
AND RELATED TEST ITEM 15

(a) Province-wide data--Inventory responses* and student performance

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers (%)	3	13	76	3	2	2	0	1
<u>Calculus</u> teachers (%)	21	16	16	0	0	0	3	43
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University teachers (%)	28	23	15	33	% correct = 62			

(b) Case study data--Inventory responses*

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers, school 1-1	0	0	1	0	0	0	0	0
<u>Calculus</u> teachers, school 1-1	0	0	0	0	0	0	0	1
<u>R and F</u> teachers, school 1-2	0	0	2	0	0	0	0	0
<u>Calculus</u> teachers, school 1-2	1	0	0	0	0	0	0	0
<u>R and F</u> teachers, school 2	0	0	1	1	0	0	0	0
<u>Calculus</u> teachers, school 2	0	1	0	0	0	0	0	0
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University 1	2	1	0	2				
University 2	1	0	1	2				

*See text, Section 2.

TABLE 4.10

DETAILED DATA ON TOPIC II-2:
RATE OF CHANGE: SLOPES, SECANTS, TANGENTS
 AND RELATED TEST ITEMS 9 AND 33

ITEM 9 FIGURES

(a) Province-wide data--Inventory responses* and student performance

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers (%)	20	13	28	2	0	1	2	34
<u>Calculus</u> teachers (%)	4	25	65	3	0	0	1	1
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University teachers (%)	10	46	23	20	% correct = 77			

(b) Case study data--Inventory responses*

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers, school 1-1	0	0	0	0	0	0	0	1
<u>Calculus</u> teachers, school 1-1	0	0	1	0	0	0	0	0
<u>R and F</u> teachers, school 1-2	1	0	0	0	0	0	0	1
<u>Calculus</u> teachers, school 1-2	0	1	0	0	0	0	0	0
<u>R and F</u> teachers, school 2	0	0	1	1	0	0	0	0
<u>Calculus</u> teachers, school 2	0	0	1	0	0	0	0	0
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University 1	0	2	0	3				
University 2	1	2	0	1				

TABLE 4.10 (continued)

ITEM 33 FIGURES

(a) Province-wide data--Inventory responses* and student performance

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers (%)	1	2	22	1	2	1	0	70
<u>Calculus</u> teachers (%)	0	0	99	0	0	0	1	0
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University teachers (%)	0	44	31	26	% correct = 49			

(b) Case study data--Inventory responses*

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers, school 1-1	0	0	0	0	0	0	0	1
<u>Calculus</u> teachers school 1-1	0	0	1	0	0	0	0	0
<u>R and F</u> teachers, school 1-2	0	0	0	0	0	0	0	2
<u>Calculus</u> teachers, school 1-2	0	0	1	0	0	0	0	0
<u>R and F</u> teachers, school 2	0	0	1	1	0	0	0	0
<u>Calculus</u> teachers, school 2	0	0	1	0	0	0	0	0
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University 1	0	2	0	3				
University 2	0	3	0	1				

*See text, Section 2.

TABLE 4.11
 DETAILED DATA ON TOPIC II-3:
DERIVATIVES OF POWERS, PRODUCTS AND QUOTIENTS
 AND RELATED TEST ITEM 13

(a) Province-wide data--Inventory responses* and student performance

	A1	A2	B	C4	C3	C2	C1	D
<u>R and F</u> teachers (%)	1	0	28	1	0	0	0	70
<u>Calculus</u> teachers (%)	0	0	100	0	0	0	0	0
	A1	A2	B	C				
University teachers (%)	5	44	23	28	% correct = 92			

(b) Case study data--Inventory responses*

	A1	A2	B	C4	C3	C2	C1	D
<u>R and F</u> teachers, school 1-1	0	0	0	0	0	0	0	1
<u>Calculus</u> teachers, school 1-1	0	0	1	0	0	0	0	0
<u>R and F</u> teachers, school 1-2	0	0	0	0	0	0	0	2
<u>Calculus</u> teachers, school 1-2	0	0	1	0	0	0	0	0
<u>R and F</u> teachers, school 2	0	1	1	0	0	0	0	0
<u>Calculus</u> teachers, school 2	0	0	1	0	0	0	0	0
	A1	A2	B	C				
University 1	1	1	0	3				
University 2	1	2	0	1				

*See text, Section 2.

TABLE 4.12

DETAILED DATA ON TOPIC II-8: MAXIMA AND MINIMA PROBLEMS
AND RELATED TEST ITEM 24

(a) Province-wide data--Inventory responses* and student performance

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F teachers (%)</u>	25	13	23	0	0	2	0	37
<u>Calculus teachers (%)</u>	3	9	87	1	0	0	0	0
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University teachers (%)	0	36	36	28	% correct = 70			

(b) Case study data--Inventory responses*

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F teachers, school 1-1</u>	0	0	0	0	0	0	0	1
<u>Calculus teachers, school 1-1</u>	0	0	1	0	0	0	0	0
<u>R and F teachers, school 1-2</u>	0	0	0	0	0	0	0	2
<u>Calculus teachers, school 1-2</u>	0	0	1	0	0	0	0	0
<u>R and F teachers, school 2</u>	0	0	2	0	0	0	0	0
<u>Calculus teachers, school 2</u>	0	0	1	0	0	0	0	0
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University 1	0	2	0	3				
University 2	0	2	1	1				

*See text, Section 2.

TABLE 4.13
DETAILED DATA ON TOPIC II-28:
DERIVATIVES OF ELEMENTARY FUNCTIONS
AND RELATED TEST ITEM 10

(a) Province-wide data--Inventory responses* and student performance

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers (%)	0	0	24	0	0	0	0	76
<u>Calculus</u> teachers (%)	0	0	100	0	0	0	0	0
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University teachers (%)	0	49	23	28	% correct = 89			

(b) Case study data--Inventory responses*

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers, school 1-1	0	0	0	0	0	0	0	0
<u>Calculus</u> teachers, school 1-1	0	0	1	0	0	0	0	0
<u>R and F</u> teachers, school 1-2	0	0	0	0	0	0	0	2
<u>Calculus</u> teachers, school 1-2	0	0	1	0	0	0	0	0
<u>R and F</u> teachers, school 2	0	0	1	0	0	0	0	1
<u>Calculus</u> teachers, school 2	0	0	1	0	0	0	0	0
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University 1	0	2	0	3				
University 2	0	3	0	1				

*See text, Section 2.

TABLE 4.14

DETAILED DATA ON PRE-GRADE THIRTEEN ITEMS 22 AND 29

ITEM 22 FIGURES

(a) Province-wide data--Inventory responses* and student performance

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers (%)	50	43	1	1	1	1	0	3
<u>Calculus</u> teachers (%)	30	61	8	0	0	2	0	0
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University teachers (%)	26	41	10	23	% correct = 53			

(b) Case study data--Inventory responses*

	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C4</u>	<u>C3</u>	<u>C2</u>	<u>C1</u>	<u>D</u>
<u>R and F</u> teachers, school 1-1	1	0	0	0	0	0	0	0
<u>Calculus</u> teachers, school 1-1	0	1	0	0	0	0	0	0
<u>R and F</u> teachers, school 1-2	1	1	0	0	0	0	0	0
<u>Calculus</u> teachers, school 1-2	0	1	0	0	0	0	0	0
<u>R and F</u> teachers, school 2	0	2	0	0	0	0	0	0
<u>Calculus</u> teachers, school 2	0	0	1	0	0	0	0	0
	<u>A1</u>	<u>A2</u>	<u>B</u>	<u>C</u>				
University 1	0	2	0	3				
University 2	3	0	0	1				

TABLE 4.14 (continued)

ITEM 29 FIGURES

(a) Province-wide data--Inventory responses* and student performance

	A1	A2	B	C4	C3	C2	C1	D
<u>R and F</u> teachers (%)	60	20	2	0	2	4	1	10
<u>Calculus</u> teachers (%)	53	29	2	0	0	3	0	14
	A1	A2	B	C				
University teachers (%)	54	18	8	20	% correct = 51			

(b) Case study data--Inventory responses*

	A1	A2	B	C4	C3	C2	C1	D
<u>R and F</u> teachers, school 1-1	1	0	0	0	0	0	0	0
<u>Calculus</u> teachers, school 1-1	1	0	0	0	0	0	0	0
<u>R and F</u> teachers, school 1-2	2	0	0	0	0	0	0	0
<u>Calculus</u> teachers, school 1-2	1	0	0	0	0	0	0	0
<u>R and F</u> teachers, school 2	0	2	0	0	0	0	0	0
<u>Calculus</u> teachers, school 2	0	0	1	0	0	0	0	0
	A1	A2	B	C				
University 1	2	0	0	3				
University 2	3	0	0	1				

*See text, Section 2.

TABLE 4.15

TOPICS WITH HIGH AND LOW ESTIMATES OF STUDENT COMPETENCE,
PROJECT II AND PROJECT III

Project III Secondary teachers			Project II % correct		Project III University teachers		Project II % correct	
Rank	Topic	Rating*		Rank	Topic	Rating*	Classification	
1	II-3	3.69	92	1	II-3	2.04	92 duplication	
2	II-28	3.67	89	2	II-28	2.00	89 duplication	
3	I-13	3.49	62	3	I-3	1.81	25-78 N.A.	
.....								
17	I-2	3.06	30	18	I-2	1.11	30 gap	
18	II-9	3.04	23	17	I-14	1.20	52 N.A.	
19	I-3	2.90	25-78	19	II-12	0.83	7 N.A.	

*Preliminary figures. See text, Section 2, for explanation and definition of scale.

APPENDIX 4A

LIST OF TOPICS INCLUDED UNDER THE HEADINGS "RELATIONS AND FUNCTIONS" AND "CALCULUS" IN THE PROJECT III CONTENT MATRIX FOR THE SECONDARY SCHOOL-UNIVERSITY INTERFACE

I. RELATIONS AND FUNCTIONS (Grade 13)

1. Function as a mapping.
2. Inverse of a function.
3. Graphs and properties of second degree relations using previously known skills.
4. Equations and graphs of conics using focus-directrix definitions.
5. Equations of conics in non-standard positions.
6. Applications.
7. Intersections of lines and conics; e.g., tangents.
8. Intersection of conics and conics.
9. Domain, range, and graph of basic trigonometric functions.
10. Standard trigonometric formulae and applications.
11. Trigonometric identity problems and equations.
12. Phase shift, period, and amplitude.
13. Translations of the plane.
14. Rotations of the plane.
15. Reflections of the plane.
16. Study of general conic.
17. (Left to be filled in with other topics covered by teacher.)

II. Calculus--Elementary Approach Topics 1-15
--Advanced Approach Topics 16-64

1. Limit of a function: intuitive approach via sequences and series.
 2. Rate of change: slopes, secants, tangents.
 3. Derivatives of powers, products, and quotients.
 4. Other derivatives: function of a function, trig functions.
 5. Applications of derivatives to tangents to curves.
 6. Further applications: velocity, acceleration.
 7. Second derivative and its use, curve-tracing.
 8. Maxima and minima problems.
 9. Rate of change problems.
 10. Differential equations; antiderivatives applied to curves and motion.
 11. Areas between curves and axes.
 12. Areas enclosed between curves.
 13. Volumes of rotation.
 14. Integration using numerical methods.
 15. Applications involving complex numbers and/or polar coordinates.
-
16. The real numbers: axioms, least upper bound, completeness.
 17. Proof by induction.
 18. Inequalities.
 19. Notation.
 20. Motivation, historical introduction.
 21. Definition and algebra of limits.
 22. Functions: definition, algebra, composition, inverse.
 23. Continuity: definition, algebra of continuous functions.

Theorems on continuous functions

24. Intermediate value.

25. Extreme value.

Derivatives

26. Definition and algebra of derivatives.

27. Chain rule.

28. Derivatives of elementary functions.

Theorems on differentiation

29. Rolle

30. Mean value

Applications of differentiation

31. Related rates.

32. Optimization.

33. Graph sketching.

34. Scientific examples.

35. l'Hôpital's Rule to limits.

Integration

36. Definition of integral and algebraic integration.

37. Fundamental Theorem of calculus.

38. Mean value theorem (MVT).

39. Application of MVT to approximation.

Techniques of integration

40. Substitution.

41. Trigonometric substitution.

42. Parts.

43. Partial fractions.

Applications of integration

44. Area.

45. Volume.

- 46. Work.
- 47. Arc length.
- 48. Improper integrals.
- 49. Taylor's theorem.
- 50. Logarithmic and exponential functions.
- 51. Hyperbolic functions.

Sequences and series

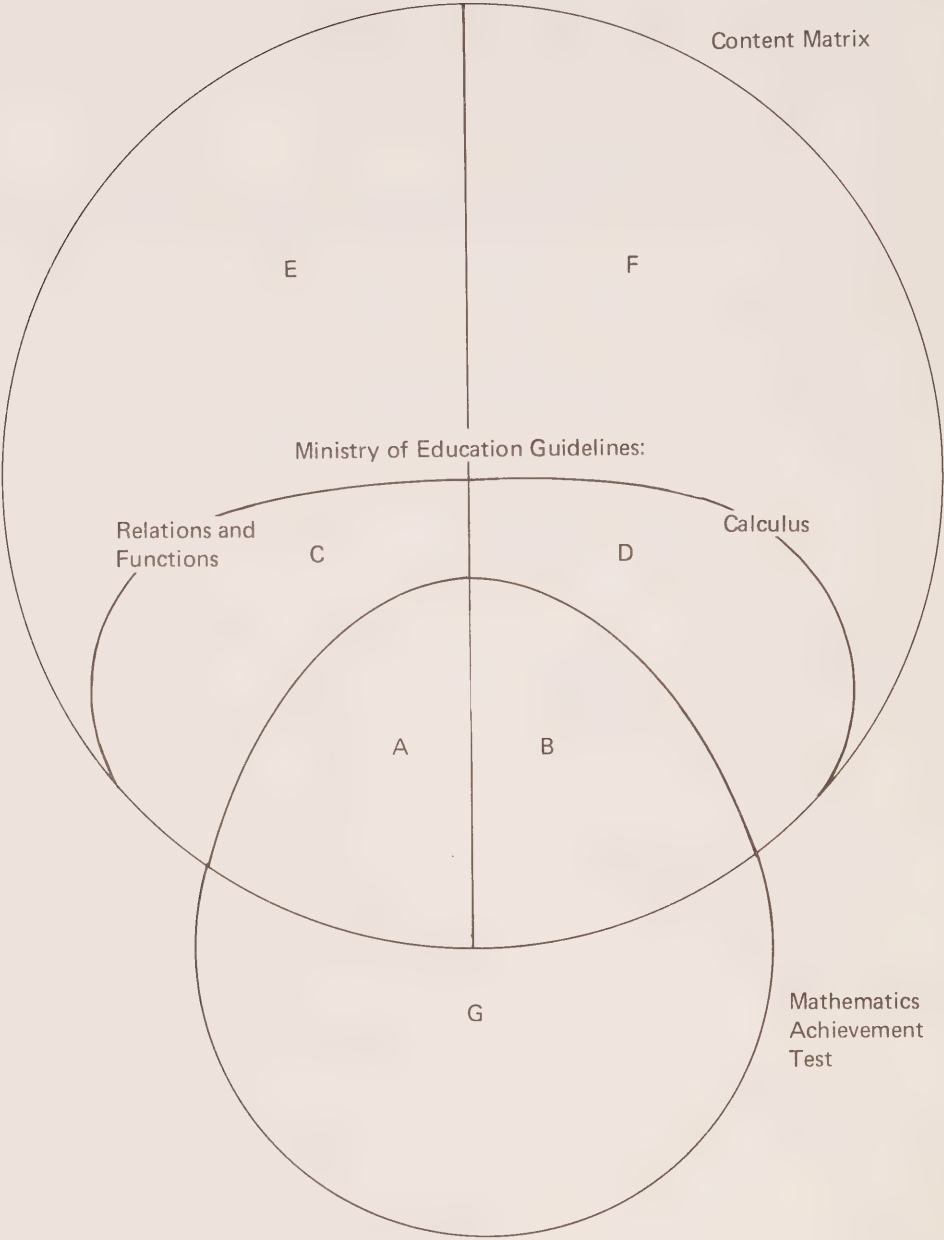
- 52. Definition and algebra of limits.
- 53. Absolute convergence.
- 54. Conditional convergence.
- 55. Basic convergence tests (e.g., ratio, root, integral, monotone).
- 56. Power series.

Elementary differential equations

- 57. Separation of variables.
- 58. General linear first order.
- 59. Partial derivatives.
- 60. Gradient.
- 61. Multiple integration.
- 62. Parametric curves.
- 63. Curvature.
- 64. Mathematical modelling.
- 65. (Left to be filled in with other topics covered by teacher.)

APPENDIX 4B

RELATIONSHIP AMONG CONTENT OF MATHEMATICS ACHIEVEMENT TEST,
CONTENT OF PROJECT III MATRIX,
AND CONTENT OF MINISTRY OF EDUCATION GUIDELINES



APPENDIX 4C
SAMPLE TOPIC TABLE

LEVEL OF COMPETENCE

		TOPIC (PROJECT III): _____					MEAN	S.D.
		0	1	2	3	4		
HIGH SCHOOL	ACTUAL:	ENTRY: _____	_____	_____	_____	_____	_____	_____
		EXIT: _____	_____	_____	_____	_____	_____	_____
	PREFERRED:	ENTRY: _____	_____	_____	_____	_____	_____	_____
		EXIT: _____	_____	_____	_____	_____	_____	_____
UNIVERSITY								
	ACTUAL:	ENTRY: _____	_____	_____	_____	_____	_____	_____
		EXIT: _____	_____	_____	_____	_____	_____	_____
	PREFERRED:	ENTRY: _____	_____	_____	_____	_____	_____	_____
		EXIT: _____	_____	_____	_____	_____	_____	_____
N (HIGH SCHOOL) = _____								
N (UNIVERSITY) = _____								

KEY: 0: NO KNOWLEDGE
 1: DESCRIPTIVE KNOWLEDGE ONLY
 2: ELEMENTARY KNOWLEDGE AND SKILL: SIMPLE APPLICATIONS ONLY
 3: WORKING GRASP: STANDARD EXERCISES AND PROBLEMS
 4: THOROUGH UNDERSTANDING, WIDE RANGE OF PROBLEMS.
 ABLE TO GENERALIZE.
 5: COMPLETE MASTERY: UNDERSTANDS THEORETICAL BASE AND
 APPLICABILITY. SOLVES NON-ROUTINE PROBLEMS.

TEST APPRAISAL INVENTORY

ITEM: _____								
HIGH SCHOOL	A1	A2	B	C4	C3	C2	C1	D
	--	--	--	--	--	--	--	--

KEY: A1: OLD KNOWLEDGE--NOT REVIEWED
 A2: OLD KNOWLEDGE--REVIEWED
 B : NEW KNOWLEDGE--ALL STUDENTS LEARN
 C4: NEW KNOWLEDGE--MORE THAN 75% STUDENTS LEARN
 C3: NEW KNOWLEDGE--51-75% STUDENTS LEARN
 C2: NEW KNOWLEDGE--26-50% STUDENTS LEARN
 C1: NEW KNOWLEDGE--1-25% STUDENTS LEARN
 D : NEW KNOWLEDGE--NO STUDENTS LEARN

UNIVERSITY:	A1	A2	B	C
	—	—	—	—

KEY: A1: OLD KNOWLEDGE--NOT REVIEWED
 A2: OLD KNOWLEDGE--REVIEWED
 B : NEW KNOWLEDGE--ALL STUDENTS EXPECTED TO LEARN
 C : OTHER

SCORE: _____% (N=1404)

CHAPTER FIVE

MATHEMATICS: THE SECONDARY SCHOOL-CAAT INTERFACE*

1. INTRODUCTION

The preceding chapter of this report gives some historical background on mathematics education in Ontario over the last twenty years, and on the events leading to a general restructuring of the mathematics program at all levels of the elementary and secondary schools.

Under the present program of studies, students at the Grade Twelve level in most schools have a number of options. Some do not study mathematics at all at this level. A small number are enrolled in "basic" courses, dealing, as the name implies, with only the most fundamental mathematical skills. Others are studying at the "general" level; their courses might be described as practically oriented, with provision for coverage of a variety of topics depending on the needs of particular classes (e.g., business arithmetic). A number are taking "advanced" courses, designed as preparation for work in mathematics at the Grade Thirteen level, and more theoretical in nature than the general courses. A small number of students in some schools have available to them "enriched" courses, similar to but more demanding than the advanced ones.

*(This chapter was prepared with the assistance of John Turner, who teaches Mathematics at Centennial Secondary School in Belleville.)

This large number of possibilities complicates enormously the task of making meaningful statements about the mathematical proficiency of Grade Twelve students as a body. Any curriculum-based testing would have to vary in content substantially from level to level, and at the general level even from class to class because of the wide number of options available to these classes. It would not be possible at all in the case of those students no longer enrolled in mathematics courses.

Another factor to be considered in choosing or developing a test appropriate to this study is, of course, the specific purpose of the study at the Grade Twelve level. What is of primary concern is the capacity of students moving from Grade Twelve into the community colleges to profitably carry out continuing work in mathematics or in subjects requiring a knowledge of basic mathematics at those colleges. Thus it is not so important to know whether students have high achievement in the content of their Grade Twelve courses as it is to know whether they have acquired those skills which they will need at the college level.

As a result of these two factors, and after lengthy discussions, the Ministries and the Project II staff agreed that the most suitable instrument for use at the Grade Twelve level would be a test which measured the competence of students in a number of basic skills which most students should have acquired by this level, and which were judged to be important to college-level mathematical success.

A new test meeting these specifications was developed for the study, and titled the Test of Arithmetic and Basic Algebra (in its French translation, the Test d'arithmétique et d'algèbre de base). Its content ranged from basic arithmetic operations with integers to the solution of quadratic equations by using the appropriate formula, but for the most part the skills tested were those which secondary teachers of mathematics and a consultant from the Ministry of Education judged that most students should have been taught in the early years of secondary school. Only

five items of the 35 on the test had content first included in the Ministry guidelines at the Grade Twelve level in either the advanced or general stream. An effort was made to avoid "trick" items, or items requiring more than the normal amount of insight or creativity, and to test the skills included in a simple and direct way.

A difficult problem arises in the interpretation of student performance on the test items, both because of lack of history on the test and because it is not a standardized test, as are, for example, those used to test both mathematics and physics at the Grade Thirteen level. The setting of a level of performance on an item which would be judged as "adequate competence" must in this case be even more arbitrary than for those tests. Taking into account the straightforward nature of the items and the basic level of their content, the investigators judged that it was reasonable to expect students to perform well. Therefore, it was decided that a success rate of 60% would be considered as indicating that students had a satisfactory grasp of the item content. A rate of 50% or more, but less than 60%, would indicate some weakness on the topic in questions, and a rate of below 50% would indicate unsatisfactory mastery of the content. For the five items on material first included in the Ministry guidelines for either the general or the advanced stream at the Grade Twelve level, these requirements would be dropped to 50% and 40% respectively, to allow for the proportion of students tested who were no longer studying mathematics or who were enrolled in a "basic" course.

2. PROJECT II DATA

Two kinds of data are available from the Project II work. First, for each item included in the test, there is a statement of the percentage of students tested who correctly answered the item. This figure has been corrected by finding such a percentage separately for each school in the study, and averaging these

percentages. Thus a figure of 53%, for example, does not mean that 53% of the students tested answered the item correctly. It does mean that in the average school 53% of the students responded correctly. In practice, the corrected figure is very close to the uncorrected one.

The second sort of data available is drawn from Test Appraisal Inventories. Teachers of both Applications of Mathematics 2 (the general stream Grade Twelve course) and Foundations of Mathematics 2 (the advanced stream course) were asked to examine the test, and to assess the content of each item as falling under one of the following headings:

A. Old knowledge that students should have on entry to the course.

A1. This knowledge is not reviewed in the course.

A2. This knowledge is reviewed in the course.

B. New knowledge that all students are expected to learn in the course.

C. New knowledge that some students are expected to learn in the course.

C1. Only 1% to 25% of students should learn this.

C2. Only 26% to 50% of students should learn this.

C3. Only 51% to 75% of students should learn this.

C4. More than 75% but not all students should learn this.

D. New knowledge that no student is expected to learn.

Teachers of various first year CAAT courses in mathematics were asked to complete a similar inventory, with the following headings:

A. Old knowledge that students should have on entry to the course.

A1. This knowledge is not reviewed in the course.

A2. This knowledge is reviewed in the course.

B. New knowledge that all students are expected to learn in the course.

C. Other.

3. PROJECT III DATA

Mathematics teachers at the Grade Twelve level and at the first year CAAT level who participated in Project III were asked to complete a Course Description Questionnaire. One portion of this questionnaire was a content matrix--a list of the topics the project staff and advisers felt might be taught or reviewed in courses at either or both of these levels. A partial list of these topics will be found in Appendix A; only those portions of the matrix covering basic arithmetic, basic algebra, quadratic functions and equations, exponential and logarithmic functions, and analytic geometry and vectors are included, since the test is concerned only with these areas. For each topic, teachers were asked to give an estimate of the level of competence of their students in handling the topic, on entry to their courses and on exit from those courses. They were also asked to state the level at which they would like their incoming students to be performing, and the level at which, given that preferred entry level, they felt they could send the student on at the end of the course. The only figures considered in this analysis will be the

reported actual secondary exit and actual CAAT entrance levels. (The figures used in this report were compiled before all data had been received. The few final figures available for Interproject Analysis suggest that changes are small.)

These estimates of competence were given on a scale of 0 to 5, with the points defined as:

- 0 Has no knowledge of this topic.
- 1 Has an awareness and descriptive knowledge of the topic only; has no ability to apply it.
- 2 Has elementary knowledge and skill, based on a simple quantitative introduction; can do simple examples and problems only.
- 3 Has a working grasp of the topic, and is able to do standard exercises and problems; has some facility in translating problems into mathematical terms, and of knowing the place of the topic in a wider context.
- 4 Has a thorough understanding of the topic and its place in mathematics, can solve a wide range of problems, prove theorems, generalize and give examples.
- 5 Has complete mastery of the topic; underlying concepts are part of his mathematical background and he can apply them to solve non-routine, symbolic, verbal, or real-life problems. He understands the theoretical basis of the topic and its limitations of applicability.

4. SUMMARY DATA

Table 5.1 presents a summary of the essential data from both projects. (As in the case of Grade Thirteen mathematics, only data from the Anglophone schools are included.) All topics included in the content matrix for which there were related test items are listed in this table, along with the mean reported actual secondary exit level and CAAT entrance levels for the topic. For each of the related test items, the table includes a summary of the Test Appraisal Inventory responses, from Applications teachers, Foundations teachers, and first year CAAT teachers, giving the percentage of each appraising the item as A, B or C (or, in the case of secondary teachers, D). As well, the table reports the percentage of students correctly answering the item.

Some comments must be made about this table. First, a number of items appear in the table more than once, because they deal with more than one of the included topics. One item is included at the end of the table which does not deal with any topic in the list (this item is on factoring integers). A large number of topics are excluded from the table because there were no relevant test items; this was to be expected given the decision to test only basic skills and to avoid specialized areas such as business arithmetic or probability.

Second, the coverage of the topics included in the list is of varying degrees. At the extremes, Topic III-2 is dealt with by seven different items, while Topic I-3, for example, is represented by only one item, which deals as well with two other topics. Clearly much more faith can be placed in conclusions based on data on a large number of items than in those drawn from a much more meagre supply of information.

One is struck on glancing at the table by the discrepancy between the level at which Grade Twelve teachers report students to be performing when they leave Grade Twelve, and the level at which CAAT teachers report receiving them. On topic after topic,

the CAAT teachers' perception of student competence is from one to one and one-half levels lower than that of the secondary teachers. It could be hypothesized that this discrepancy is due to the probable continuation into Grade Thirteen of the more able Grade Twelve students, so that those entering the CAATs are on the average below the norm of performance for students in that grade. However, the discrepancy is found in the data for all other subjects considered in the study as well, at both the Grade Twelve and Grade Thirteen levels, and is in some cases more sizeable than it is here. A more plausible suggestion is that secondary teachers and CAAT teachers have perceived not the students, but the scale, differently--that what a secondary teacher considers to be a working grasp of a topic, for instance, is considered by a CAAT teacher to be elementary quantitative knowledge. In any case, the discrepancy makes the figures as they stand meaningless.

What is interesting, though, is the correlation between the two sets of figures. If they are used to rank-order the topics, there is considerable agreement between the two groups as to which topics the students are most and least capable in dealing with. And both sets of ratings correlate significantly with student performance on the items related to the topics, with the correspondence being slightly closer in the case of the secondary teachers. This degree of agreement between the data from the two projects provides some confirmation of the reliability of each set of data.

Table 5.1 provides a setting for examining the data from the two projects to see whether there is evidence of the existence of gaps or duplications. A gap will be said to exist when CAAT teachers have the expectation that incoming students have previous knowledge of a topic or of the content of a particular item, but the student performance data indicate a low level of such knowledge. A duplication occurs when CAAT teachers are spending time teaching content that students already know well. The criteria have not been rigidly set, but in rough terms a student success rate of 50% to 60% on an item whose content is

treated as previous knowledge by at least 40% of CAAT teachers defines a minor gap. If the success rate is below 50% on such an item, the gap is defined as major. A duplication is roughly defined as the case where the success rate on an item is 60% or more, but 30% or more of CAAT teachers treat its content as new knowledge. These guidelines are subject to modification by the Project III data on reported exit and entrance levels in the topic, and by peculiarities of particular items. Data gathered from Grade Twelve teachers will not be used in identifying a gap or duplication, but will be examined for indications of the causes of gaps and duplications.

The next section will examine in turn each topic listed in Table 5.1, along with its related items.

5. INDIVIDUAL TOPICS

5.1 General Topic I: Basic Arithmetic

Both secondary teachers and CAAT teachers rated student competency in Topic I-1, "Fundamental arithmetic operations with fractions, decimals and integers", very high--first on the list for CAAT teachers, and third for secondary teachers, out of the 17 topics examined here. CAAT teachers overwhelmingly expected students to have knowledge of the content of the six related items when they arrived in their courses; for four of the six items 100% of these teachers assigned an A rating, and for both of the others the figure was over 90%. Student performance was well over the 60% criterion on five of the items, dropping to 55% (in the "minor gap" range) on one item. Although a closer look will be taken at this item in the next section of this chapter, the gap problem with this topic is at most minor, and limited in extent. There is no duplication evident.

Competency in Topic I-3, "Percentage", was also ranked high by both groups of teachers; secondary and CAAT teachers placed it fourth and second respectively. 86% of CAAT teachers assumed that incoming students would have previous knowledge of the content of the one test item dealing with percentage. However, only 31% of students answered the item correctly. This item deserves close scrutiny as strongly indicating a gap.

Topic I-5 is "Measurement: use of exact and approximate numbers (error, precision, accuracy, rounding off)". Secondary teachers rated student competency in this topic very low on their scale, in fifteenth place, and CAAT teachers placed it last. However, both secondary teachers and CAAT teachers felt students should know the content of the related test item (the same item as in the previous topic), although only 31% of the students succeeded in giving the correct answer. A gap may well be indicated here.

In the case of Topic I-6, "Scientific notation: conversion to and from", secondary and CAAT teachers disagreed about the level of student knowledge. Secondary teachers ranked the topic seventh, but CAAT teachers put it much lower, in twelfth place. Slightly fewer CAAT teachers than on previous topics expected incoming students to have been taught the content of the related item, but the number was still sizeable, at 78%. Student performance was poor, with only 38% giving the correct answer to the item. Again a gap is indicated.

The same item dealt with Topic I-7, "Scientific notation: use in computation and estimation". This topic is closely related to Topic I-6, but both sets of teachers placed it somewhat lower than that topic on the competency scale--eleventh and fifteenth for secondary and CAAT teachers respectively. As stated above, the content was expected on entry by 78% of CAAT teachers, but only 39% of students correctly answered the item.

The first impression of the summary data on General Topic I is that students are dealing well with fundamental arithmetic operations. There are indications of problems with percentage and with scientific notation, where student test performance was low although most CAAT teachers expected students to have mastered these skills before entering their courses.

5.2 General Topic III: Basic Algebra

This section of the matrix has a complex interrelationship with the test. A number of topics are related to several items, and a number of items deal with several topics. Item 19, for example, is included in Table 5.1 under Topic III-2, Topic III-9, Topic III-10 and Topic III-14. A little more caution must be used in interpreting this section than the others discussed because of these overlaps.

There are seven test items dealing with Topic III-2, "Manipulating, simplifying and evaluating algebraic expressions". Secondary and CAAT teachers ranked student competency high--second and fourth respectively. CAAT expectations that students would know the content of the items were high; the percentage considering the content to be old knowledge ranged from 68 to 89. Student performance was generally good on these items, exceeding the 60% success mark on five of the seven items. On the remaining two the percentages answering correctly were 58 and 50, and these two items will be examined for minor gap problems.

A difference in perception appeared in the case of Topic III-3, "Linear equations and word problems involving one unknown", with secondary teachers placing student competency at the midpoint of their scale and CAAT teachers rating it third highest. The two related test items show the same inconsistency; performance on one was poor (31% correct), while the other suggested at most a minor gap (57% correct). The item content was in both cases treated as old knowledge by over 80% of CAAT teachers.

Only one item covered Topic III-4, "Systems of linear equations in 2 unknowns and applications". CAAT teachers rated student knowledge of this topic somewhat higher, at sixth, than did secondary teachers, who ranked it tenth. The item content was regarded as previous knowledge by 68% of CAAT teachers, and student performance suggests a minor gap, with 57% choosing the correct response.

Topic III-9, "Factoring: various types, complex fractions", was well up on the list in the competency ratings of secondary teachers, in sixth place. CAAT teachers placed it ninth, at the midpoint of the scale. The percentage of CAAT teachers expecting student knowledge of the content of each of the four related test items ranged from 57% to 68%. Performance on two of the items was well above the 60% level. The success rates of 50% and 48% on the other two items indicate the possibility of a partial gap in the topic.

Both groups of teachers rated Topic III-10, "Operations with rational algebraic fractions", eleventh on their scales of competency. From 60% to 70% of CAAT teachers expected previous knowledge of the content of the four related test items. One item was well done, but results on two others suggest a minor gap, and the fourth item was poorly enough answered that its content must tentatively be considered as indicating a major gap.

Two items dealt with Topic III-12, "Operations with radicals and irrationals". Secondary teachers placed student knowledge of the topic just above the midpoint of their rank-ordering, but CAAT teachers rated it near the bottom, in fifteenth place. The content of the related test items was expected on entry by considerably fewer CAAT teachers than was the case for earlier topics, but the number expecting it was still substantial--54% for each of the two items. Students performed well (65% correct) on one item, but only 31% correctly answered the second item. A gap is indicated here.

The situation is better with Topic III-14, "Four fundamental operations on polynomials", which is the topic in which secondary teachers rated student competence highest. CAAT teachers placed it seventh on their list. The content of the five related test items was expected at CAAT entry by large numbers of teachers at that level--from 65% to 89%, depending on the item. Students surpassed the 60% performance criterion on three of the five items; two, however, indicate a minor gap with success rates of 58% and 50% respectively.

Secondary teacher expectations of competency were low for Topic III-15, "Manipulation, rearrangement, evaluation of algebraic formulae", which was ranked fourteenth. CAAT teachers placed the topic at the midpoint of the scale, in ninth place. The one related test item suggests that these low expectations are justified. Although 78% of CAAT teachers treated the item content as previously learned, only 41% of students chose the correct response.

Altogether there are nineteen test items dealing with basic algebra (as defined in the content matrix). The percentage of CAAT teachers expecting previous knowledge of the item content ranged from 54% to 89%--a sizeable figure even at the lower boundary. Student performance was variable, with success on five items falling below 50%, success on four more items falling between 50% and 60%, and more than 60% of students correctly answering each of the remaining nine. Thus one-half of the items considered under the heading of basic algebra suggest gaps of varying degrees of seriousness.

5.3 General Topic IV: Quadratic Functions and Equations

Only one topic in this section of the matrix is represented by test items; this is Topic IV-7, "Quadratic equations: completing the square, formula, problems". CAAT teachers had a higher degree of confidence in student knowledge of this topic than did secondary teachers, ranking it seventh as opposed to the secondary

ranking of thirteenth. Knowledge of the content of the two test items dealing with quadratic equations was expected on entry by better than half the CAAT teachers. The student success rates of 48% and 25% indicate the existence of a gap.

5.4 General Topic V: Exponential and Logarithmic Functions

Again only one topic under this heading was tested--Topic V-1, "Exponents: whole numbers, integers, rationals". Both groups of teachers placed this topic fifth on the rank-ordered list of student competency. Each of the four test items had content considered previous knowledge by 70% or more of CAAT teachers. Overall student performance was somewhat discouraging, however. Only one item was successfully completed by 60% or more of the students; one fell into the "minor gap" category; and two suggested serious gaps.

5.5 General Topic VII: Analytic Geometry and Vectors

Topic VII-2, "Derivation of various forms of the equation (of a straight line): two points, slope and point, intercepts, etc.", was treated by three test items. This topic was ranked last by secondary teachers, and thirteenth by CAAT teachers, in terms of student competency. The content of each of the three items was retaught at the CAAT level by 30% or more of teachers. Although student performance was good on one item, the other two showed poor performance. The first of these three items may indicate some duplication in dealing with this topic.

Topic VII-3, "Identifying, constructing and graphing a straight line from given data", was also placed very low on both sets of ratings of student knowledge--sixteenth by secondary teachers and fourteenth by CAAT teachers. Student performance on the related test item supports these assessments; only 39% gave the correct response, although well over half the CAAT teachers

considered the content to be old knowledge at that level. A gap seems to exist here.

5.6 Item Not Related to Content Matrix Topics

One test item cannot easily be classified under any of the topic headings in the matrix. This item (number 2) requires students to find the greatest common factor of two integers. Although CAAT teachers unanimously expected students to have this skill on their arrival at the colleges, only 42% of students were able to give the correct response to the item.

6. DUPLICATIONS

The only topic for which the question of duplication has been raised in the previous section is Topic VII-2, "Derivations of various forms of the equation (of a straight line): two points, slope and point, intercepts, etc." Three items, numbers 28, 29 and 30, test this topic, and, as may be seen from Table 5.1, CAAT treatment of the item content is very similar for each case. The question of duplication arises only for item 28, because, of the three items, it is the only one on which students performed well, thus casting doubt on the need for the widespread reteaching of the material at the college level.

There seems no doubt that there is some duplication in the teaching of this particular item content, especially when it is considered that a more detailed breakdown of inventory responses would show that, of the 38% of CAAT teachers who considered this old knowledge, 35% spent time reviewing it. Thus altogether, of the 68% of teachers finding this knowledge relevant to their college courses, 65% spent time either on reviewing the material or on teaching it.

What muddies this conclusion, however, is the large discrepancy between student performance on this item and on item 30. The well-done item presents students with a straight-line equation in standard form and asks them to state the slope and the y-intercept. The poorly-done item gives them the y-intercept and some information about the slope, and asks for the equation. The same skills are involved in each item, but item 30 requires a slightly more complex application of the skills which the students were unable to perform. It may well be, then, that the judgement of the CAAT teachers has been that since it is clearly necessary to reteach the slope-point form of the equation at a level just above basic, it is wise to first make certain that students know the basics, through review or formal reteaching. In any case, the conflicting evidence of item 30 strongly suggests that the material is known only at the introductory level, and that the duplication indicated here is in many respects illusory.

7. GAPS

Gaps, on the other hand, have been at least suggested in a number of the topics included in Table 5.1. We shall look at each of these topics in turn, examining the key items in more detail than in the previous section.

7.1 General Topic I: Basic Arithmetic

The only item under Topic I-1 which needs examination is item 3, which produced test results in the "minor gap" category. It required students to realize that, given a fraction preceded by a negative sign and with a negative denominator, one could remove both negative signs. Detailed data indicate that 70% of Applications teachers and 90% of Foundations teachers assumed this knowledge without review at the beginning of Grade Twelve courses. In view of the unsatisfactory test results, these teachers might be well advised to make certain that their

students are adept in this skill by providing some short time for reviewing it during the Grade Twelve course.

Topics I-3 and I-5 are each represented by only one item, the same in each case; the item also deals with Topic III-3. It is risky to draw any conclusions under these circumstances about where the deficiency lies which produced the poor results. However, the item was structured in such a way that the most probable "rounding-off" errors were not included as possible responses, and the most probable errors resulting from students not being sure how to tackle the problem were included. Students were asked to find, correct to one decimal place, an unknown number, given the value of a certain percentage of that number. To judge by the response pattern, many students interpreted the problem in reverse, and found the specified percentage of the given number, either correctly (21%) or with a misplaced decimal in their answer (20%). Thus the poor results may with some confidence be attributed largely to lack of facility in handling percentage (Topic I-3); any students weak in rounding off (Topic I-5) would not have found the result of their error as a possible choice, and would be included in the relatively small percentage of students not answering the item. It is in the realm of percentage, then, that the gap exists. Detailed data on the item again show that many Grade Twelve teachers did not review this content in their classes; 43% of Applications teachers and 74% of Foundations teachers simply took it for granted. As in the previous case, some time for review is indicated.

The next problem item is again the sole representative of two different topics, although these are much more closely related than in the preceding case in that they cover two different aspects of scientific notation. The item in question, item 8, presents students with a calculation to be carried out using numbers given in ordinary rather than scientific notation, but requires the answer to be chosen from among five possibilities written in scientific notation. The numbers involved are simple except for the presence of zeroes (e.g., 0.0006), and the only difference among the answers is that differing powers of

10 are given. Thus the question is contaminated as little as possible by other knowledge. 35% of both Applications and Foundations teachers were confident enough that students could handle this material on entering their courses to assume it without review. CAAT teachers were, probably wisely, a little less confident; although 78% expected students to have met the content before, only 14% neglected to at least review it. The content appears to be important at the CAAT level, since almost all teachers either reviewed or taught it, and again Grade Twelve teachers could profitably give it some review time.

7.2 General Topic III: Basic Algebra

There are only two problem items on Topic III-2, and both indicate minor gaps. The content of item 15 was assumed without review by 19% of CAAT teachers, and no more than reviewed by another 70%. It was considered old knowledge on entry by 95% of Applications teachers and all Foundations teachers. The item is very simple, requiring students to add two binomials. An examination of the response pattern indicates that by far the most common error (30% of students) was a careless misreading of the item which resulted in the student multiplying these binomials rather than adding them. The students demonstrated in other items that they could perform more complex operations on algebraic expressions with adequate skill, and the results on this item must be regarded as a fluke stemming from carelessness rather than from lack of competence in the skill. No gap can be deduced from the poor performance in this case. Item 19, however, presents a different picture. Students were asked to simplify an algebraic fraction with two terms in the numerator and one in the denominator which was a factor of both the numerator terms. Although students performed well on the preceding item, which was similar except that the numerator had only one term, they were poor in dealing with the more complex problem. The most common error (23% of students) was to divide only one of the terms in the numerator by the denominator. This skill was widely reviewed in Grade Twelve courses (72% of

Applications teachers and 56% of Foundations teachers) or taught as new material in the Applications course (14% of teachers). but student competence was still poor. CAAT teachers are apparently justified in reviewing the material (62%) or reteaching it (16%).

Both items dealing with Topic III-3 suggested gaps. Item 9 elicited very poor performance; however, this is the same item discussed above under Topics I-3 and I-5. As was stated there, the main difficulty seemed to be dealing with percentage, and little can be decided from its results about student capability in the present topic. Item 23 should be a better indicator, since it does not deal with any other topic, asking the student to solve a single linear equation in one unknown. The only complication consists in the need for the student to first collect like terms. Performance was near the 60% criterion, but again CAAT teachers are wise to review the skill; of the 81% considering it old knowledge, 70% gave it some review. It was also reviewed by Applications teachers at the earlier level by 62% of the 95% treating it as former knowledge, but apparently not with sufficient effect. A minor gap continues to exist.

Item 24 is the only item testing Topic III-4. Here too performance was almost adequate, but lower than should be expected on an item whose content was assumed by the large majority of teachers of both Grade Twelve courses to have been learned in earlier years, and was widely reviewed in both courses (62% of Applications teachers and 64% of Foundations teachers spent time reviewing this skill). No CAAT teacher who responded to the item assumed that students could handle the solution of a system of linear equations in two unknowns without review--a reasonable belief in view of the minor gap revealed here.

Of the four items on Topic III-9, two indicate the existence of a gap. The two problem items are the one discussed under Topic III-2 (simplifying an algebraic fraction) and one requiring the solution of a quadratic equation by factoring. The former has already been discussed above, and appears to indicate a gap in this topic as well. The latter is a skill often not taught

until the Grade Twelve level, and somewhat poorer performance is to be expected because many students are no longer studying mathematics at this level. Most errors (19% of students) involved changing the signs of the roots; it may well be that students factored the quadratic correctly, but became confused on the next step because they had not been formally taught the skill. However, this must still be classified as a gap because of perhaps unreasonable CAAT expectations of previous knowledge (although all CAAT teachers dealing with the skill did at least review it).

The four items on Topic III-10 must be examined here. The well-done item and one of the poorly-done ones (item 19) have already been discussed under Topic III-2; they involve the simplification of rational algebraic fractions where the denominator is a factor of the numerator. This skill was well handled when the numerator consisted of a single term, but poorly handled when there were two terms in the numerator. The most poorly done item (item 34) required students to add two rational algebraic expressions with unlike denominators. 42% of the students apparently simply added numerators and added denominators; this is considerably above the 33% who did the item correctly. This skill was not as widely expected or taught in Applications courses as were many of the others on the test, with 15% of these teachers omitting coverage with some or all of their students. Those Applications teachers who did deal with the material for the most part either reviewed or taught it; only a few expected it without review. About two-thirds of Foundations teachers dealing with the material reviewed it, about one-third assumed it without review. Clearly college teachers are wise to review or reteach this skill; almost none expected student competence without review, and a substantial number taught it as new. Item 35 required the most manipulation of the items on this topic; students were asked to simplify a rational algebraic fraction whose numerator and denominator each contained a further fraction. Extensive review in both Grade Twelve courses had apparently had some result here, with competence approaching the adequate level. However, general review (58%) or reteaching (19%)

at the CAAT level is probably justified by the "minor gap" picture. Overall, one major gap and two minor gaps are confirmed in this topic.

Under Topic III-12, students performed satisfactorily in adding two radicals. They did very badly, however, at rationalizing the denominator of a fraction with radical numerator and denominator, with the most frequent response (35% of students) being to simply remove the radical signs. This was not considered a new skill by most Grade Twelve teachers of either stream, though it was reviewed by a large percentage of them. This review apparently did not have the desired effect; most CAAT teachers dealing with the skill found further review or reteaching necessary.

The two problem items on Topic III-14 both suggest minor gaps. One is the often-mentioned item 19, discussed above under Topic III-2. The other has also been examined in connection with that topic, and poor performance on it dismissed as proceeding from causes other than lack of competence in the required skill. We are left with the minor gap suggested by item 19.

The one item on Topic III-15 was treated as old knowledge by Grade Twelve teachers in both streams, not even reviewed by about one-quarter of Applications teachers and about one-half of Foundations teachers. The poor results on the item confirm the CAAT practice of reviewing the material, as did 76% of the 78% treating it as previous knowledge. It is unfortunate, here as elsewhere, that time needs to be spent in reviewing badly learned material rather than in proceeding to more complex work.

7.3 General Topic IV: Quadratic Functions and Equations

Performance was poor on both items dealing with quadratic equations. One item admitted of a solution by factoring; the second required application of the appropriate formula, and was much more poorly done. The first item has been discussed under

Topic III-9. 82% of general stream students would not be taught the factoring method until Grade Twelve if at all; the same holds for 50% of advanced stream students. The use of a formula was expected on Grade Twelve entry by only 8% of Applications teachers and the same number of Foundations teachers. The gap here may well be a function of too high expectations at the CAAT level of students who have not covered the material.

7.4 General Topic V: Exponential and Logarithmic Functions

Four items asked students to deal with exponents. The only one on which performance was good was an item involving numerical substitution into an algebraic expression which included an exponent. Items 10, 11 and 12 dealt respectively with a negative integral exponent, a power of a power, and fractional exponents, all in numerical rather than algebraic terms. The usual error in item 10, which asked students to evaluate a fraction with a negative integral exponent, was to treat the exponent as positive and insert a negative sign before the result. Errors were varied in the "power of a power" item (number 11). Item 12 required students to divide an integer with a fractional exponent by the same integer with a different fractional exponent. 23% of the students chose to divide the first exponent by the second, leaving the base unchanged. In spite of suggested earlier treatment in the Ministry guidelines, a large number of students apparently are not taught much about exponents until the Grade Twelve level. The percentages of teachers expecting incoming students to have covered the content of these three items were respectively: in Applications, 60, 68 and 44; in Foundations, 58, 59 and 38. So again it may be unrealistic for CAAT teachers to expect as much previous knowledge as they did in this area (although they were more realistic in almost without exception treating this as review material when they did expect it to have been covered).

7.5 General Topic VII: Analytic Geometry and Vectors

Topic VII-3 appears to suffer from a minor gap, with the success rate on item 27 being only 39% in spite of the 57% of CAAT teachers expecting previous knowledge of the topic. However, 54% of these teachers do review the material. The greatest number of errors on the item were errors of omission; 24% of the students did not attempt to answer it. It required students to draw a straight line graph, given the equation of the line in standard form. Both Applications and Foundations teachers described it as pre-Grade Twelve work, though widely reviewed at that level. This appears to be a major gap.

7.6 Item Not Related to Content Matrix Topics

Item 2 required students to find the highest common factor of two integers (although the words "highest common factor" were not used). Just under half the students chose a response which was a factor of both integers, but not the highest such number. All CAAT teachers expected students to have this skill, and only 30% reviewed it. The great majority of both Applications and Foundations teachers considered it previous knowledge; almost none reviewed it. Factoring integers would clearly profit from some attention at the Grade Twelve level, considering the low success rate on the item.

8. SUMMARY

In general, the performance of students on this test supports the usual complaint of CAAT teachers that their incoming students do not possess the degree of skill they should have in doing basic arithmetic and algebra. In certain cases, teachers at the college level appeared to expect more than a reasonable amount of their incoming students, many of whom would not have studied mathematics at the Grade Twelve level (although it should be

pointed out that Grade Twelve mathematics is a prerequisite for technical courses at the CAATs, and teachers of such courses are justified in expecting more knowledge of their students). But much more often these students appeared not to have mastered material taught to them in the earlier years of secondary school, when all or almost all of them were still enrolled in mathematics courses. The situation is probably more serious than this analysis indicates, since the sample of students included many who planned to continue their secondary studies, presumably with the aim of eventually attending university. It seems safe to assume that these would be on the average the more able students, so that the performance of those students making the move from Grade Twelve into the CAATs would be below the overall performance reported here. This assumption is, in fact, confirmed in the Project II report, which presents comparative score distributions for various subgroups of students.

Table 5.2 presents a rough categorization by topic of those gaps which appear to be real. It makes little sense to total the figures in the various columns because of the frequent appearance of an item under more than one topic. However, the general impression given by the table is discouraging. There is no topic for which all related items were answered satisfactorily by the students, although the situation approaches the satisfactory for a few. The items falling into the column heading "No. showing poor performance, no gap" should not be interpreted as indicating that no gap exists; the most that can be said is that the poor performance can be explained in other ways, so that the question of the existence of a gap appears unresolved. In three topics, some or all gaps are attributed to overly high expectations at the CAAT level that students will have covered material widely taught for the first time in Grade Twelve; our assumption is that a significant number of CAAT entrants will not have continued their mathematical studies to this level (except, as noted above, in the case of technical courses). In most topics, the problems can confidently be assessed as the result of inadequate learning in the earlier grades of secondary school. It is disturbing that in many of these cases the skills have not been given adequate

attention in the form of review in Grade Twelve in spite of the clear deficiencies in student knowledge. (It is of interest in this regard that most of the gaps found at the Grade Thirteen level in mathematics and discussed in the previous chapter could also be traced back to previous years, where material was apparently not properly mastered; again, the teachers of the courses in the final secondary year assumed that students know this material and often did not review it.)

Remedial action could be taken at all three levels. In the earlier secondary years, more rigour is required to make sure that students have some understanding of how to apply the skills they are taught and which they will need at higher levels. Grade Twelve teachers should spend more time reviewing these skills and assuring that students can use them competently. CAAT teachers should resign themselves to the fact that incoming students will have an imperfect knowledge of exponents and quadratic equations in particular; if the students are known to be competent in the more basic areas, these teachers will require much less time for review and can use the time saved to cover these two problem areas and to proceed to more difficult work. (An alternative, of course, is to require all secondary students to carry their enrolment in mathematics courses to the Grade Twelve level; presumably they would then be skilled in exponents and quadratic equations as well, and would provide the CAAT teachers with even more time for new work.)

There is a further question, peripheral to this report but perhaps worthy of brief mention. Many students entering community colleges do so, not from the Grade Twelve high school level, but from the work force after a period of absence from school. A number of these entrants have had very little formal secondary education, and changes in emphasis in the secondary school curriculum would not affect their generally low standard of performance in mathematics. It might thus be claimed that there is little to gain in upgrading the skills of students staying in secondary school until Grade Twelve, since the CAATs will still need to spend as much time as they do now reviewing and

reteaching topics which these other students have not met. However, it is plausible at least that, given a consistent standard of performance from Grade Twelve students, the CAATs could organize a remedial program designed to bring other entrants up to that standard before admission to regular mathematics courses or courses requiring this mathematical knowledge.

There is one optimistic note that can be struck here. At the same time that complaints have been increasing about the lack of mathematical knowledge of secondary students, secondary enrolment, particularly in the senior grades, has been rising sharply (see Chapter 1 of this report). Many of the students now in Grade Twelve are students who, had they entered secondary school ten or fifteen years ago, would have been very unlikely to remain in school for four years. Consequently, comparisons of the performance of the "average" student of today with his counterpart of that period can be highly misleading, since the earlier group of students was much more highly selected.

The best students tested performed very well indeed on this test, and the project staff have seen no evidence that leads them to believe that these students know less than the best students of those earlier days. What has changed is the composition of the Grade Twelve classes in almost every school. One group of students is not being shortchanged; rather, another group is getting the opportunity to acquire (albeit imperfectly) skills that formerly would have been out of their reach.

TABLE 5.1

SSGD MATHEMATICS: SUMMARY OF ESSENTIAL DATA

Topic	Reported secondary exit level* level*	Reported CAAT entrance level*	Item No.	Assessment by**												% correct
				Applications teachers				Foundations teachers				CAAT teachers				
				A	B	C	D	A	B	C	D	A	B	C		
I-1	3.17	2.23	1	99	0	1	0	92	8	0	0	100	0	0	71	
			3	70	30	0	0	90	10	0	0	92	3	5	55	
			4	100	0	0	0	100	0	0	0	100	0	0	70	
			5	100	0	0	0	100	0	0	0	100	0	0	70	
			6	100	0	0	0	100	0	0	0	97	3	0	72	
			7	100	0	0	0	100	0	0	0	100	0	0	76	
I-3	3.10	1.97	9	90	5	3	3	95	2	1	2	86	8	5	31	
I-5	2.48	1.06	9	90	5	3	3	95	2	1	2	86	8	5	31	
I-6	2.97	1.43	8	88	11	1	0	89	9	2	0	78	14	8	38	
I-7	2.77	1.23	8	88	11	1	0	89	9	2	0	78	14	8	38	
III-2	3.18	1.79	13	92	5	1	1	99	1	0	0	84	11	5	68	
			14	84	13	3	0	96	4	0	0	84	11	5	72	
			15	95	3	3	0	100	0	0	0	89	0	11	58	
			16	94	5	1	0	100	0	0	0	76	14	11	88	
			17	89	7	4	0	100	0	0	0	70	16	14	69	
			18	88	9	3	0	97	2	1	0	70	16	14	63	
			19	83	14	3	0	98	2	0	0	68	16	16	50	

TABLE 5.1 (continued)

Topic	Reported secondary exit level* level*	Reported CAAT entrance level*	Item No.	Assessment by**												% correct
				Applications teachers				Foundations teachers				CAAT teachers				
				A	B	C	D	A	B	C	D	A	B	C		
III-3	2.84	1.80	9	90	5	3	3	95	2	1	2	86	8	5	31	
			23	95	4	0	1	98	2	0	0	81	14	5	57	
III-4	2.78	1.67	24	88	9	1	1	96	4	0	0	68	16	16	57	
III-9	3.02	1.54	19	83	14	3	0	98	2	0	0	68	16	16	50	
			20	71	24	4	1	98	2	0	0	65	16	19	68	
			22	68	29	3	0	95	5	0	0	70	11	19	69	
			25	18	66	11	5	50	47	3	0	57	22	22	48	
III-10	2.77	1.44	18	88	9	3	0	97	2	1	0	70	16	14	63	
			19	83	14	3	0	98	2	0	0	68	16	16	50	
			34	50	34	10	5	95	4	0	1	60	22	19	33	
			35	55	22	14	10	94	3	1	2	61	19	19	52	
III-12	2.94	1.23	32	78	16	4	3	96	4	0	0	54	22	24	65	
			33	72	18	5	4	95	5	0	0	54	19	27	31	
III-14	3.20	1.56	15	95	3	3	0	100	0	0	0	89	0	11	58	
			16	94	5	1	0	100	0	0	0	76	14	11	88	
			17	89	7	4	0	100	0	0	0	70	16	14	68	
			19	83	14	3	0	98	2	0	0	68	16	16	50	
			21	88	9	3	0	100	0	0	0	65	16	19	79	
III-15	2.68	1.54	31	81	15	3	1	98	2	0	0	78	8	14	41	

TABLE 5.1 (continued)

Topic	Reported secondary exit level*	Reported CAAT entrance level*	Item No.	Assessment by**												% correct		
				Applications teachers				Foundations teachers				CAAT teachers						
				A		B		C		D		A		B			C	
				A	B	A	B	A	B	A	B	A	B	A	B			
IV-7	2.76	1.56	25	18	66	11	5	50	47	3	0	57	22	22	48			
			26	8	75	12	5	8	88	4	0	54	24	22	25			
V-1	3.05	1.69	10	60	27	9	4	58	39	3	0	70	19	11	37			
			11	68	21	9	1	59	39	2	0	73	16	11	53			
			12	44	39	13	4	38	58	4	0	70	19	11	44			
			14	84	13	3	0	96	4	0	0	84	11	5	72			
VII-2	2.00	1.41	28	82	11	3	4	95	5	0	0	38	30	32	66			
			29	87	5	3	5	96	4	0	0	32	43	24	42			
			30	80	5	6	10	94	6	0	0	32	32	35	46			
VII-3	2.27	1.32	27	87	8	3	3	94	6	0	0	57	30	14	39			
-----			2	99	0	1	0	98	0	1	1	100	0	0	42			

*Preliminary figures. See text, Section 3, for explanation and scale.

**See text, Section 2, for explanation of assessments.

TABLE 5.2
DISTRIBUTION OF GAPS

Topic	Total no. of items	No. showing good performance, no duplication	No. showing poor performance				
			no gap expectations	minor gap due to high CAAT	minor gap due to inadequate learning	major gap due to high CAAT expectations	major gap due to inadequate learning
I-1	6	5	0	0	1	0	0
I-3	1	0	0	0	0	0	1
I-5	1	0	1	0	0	0	0
I-6	1	0	0	0	0	0	1
I-7	1	0	0	0	0	0	1
III-2	7	5	1	0	1	0	0
III-3	2	0	1	0	1	0	0
III-4	1	0	0	0	1	0	0
III-9	4	2	0	0	1	1	0
III-10	4	1	0	0	2	0	1
III-12	2	1	0	0	0	0	1
III-14	5	3	1	0	1	0	0
III-15	1	0	0	0	0	0	1
IV-7	2	0	0	0	0	2	0
V-1	4	1	0	1	0	2	0
VII-2	3	1	2	0	0	0	0
VII-3	1	0	0	0	0	0	1
--	1	0	0	0	0	0	1

CHAPTER SIX

PHYSICS*

1. INTRODUCTION

There are different sorts of "gaps" and "duplications" possible in the content of consecutive courses in a subject. A gap may occur when teachers at the lower level are not covering, or not covering in sufficient depth, material which teachers at the higher level are assuming the students know. It may also occur when teachers at the lower level are covering the material, but for some reason the students are not learning it. A duplication--the reteaching of material the students already know--is usually the result of the same material being covered in both courses. It can also happen that the material has not been included in the course at the lower level, but that students have learned it in some other context before their entry to the higher level. This is somewhat less likely in an area such as physics than in, for example, history or English literature, but it is theoretically possible.

In order, then, to discover whether gaps and duplications exist between secondary school physics courses and those in universities, it is necessary to examine not only what teachers teach, but also what students learn. What appears to be a gap between the two programs may on closer examination prove otherwise, as the students may have acquired the untaught

*(This chapter was written with the assistance of Douglas R. Scott, Head of the Physics Department, Downsview Secondary School, North York.)

material independently. Similarly, an apparent duplication may be a necessary reteaching of material that was taught but not learned at the lower level. Thus, much more information can be gained by combining the data from Project II on student performance with the data from Project III on program content at both levels than could be gained from the separated data.

A crucial question about the data from both projects is: How much is enough? How many students need to answer a test question correctly for us to be able to say that students have a sufficient grasp of the item content? How many teachers need to agree that students should be able to answer the item to make our expectation of a high success rate reasonable? How many teachers need to tell us that they cover a particular topic in depth so that we can state that as a general pattern? The decision in each case must be highly arbitrary, in the absence of longitudinal data. The first question, for instance, depends strongly on the nature and difficulty of the test employed. In this study, the test used was of high difficulty. It was originally designed as an instrument to be used as a criterion for university admission, and to discriminate more highly between "good" and "very good" than between "good" and "bad". It would be unreasonable, therefore, to expect the average SSHGD level student to achieve a very high score. On the basis of results from the previous use of the test, the investigators and their advisers made the judgement that a 60% success rate on an item indicated that the students had very good knowledge of the content. 50% was considered good knowledge, 30% adequate knowledge, 15% poor knowledge, and below 15% very poor knowledge.

It was not possible to apply a similar simple and arbitrary judgement to teacher assessments of whether their students should have covered the content of a particular item, either before or during their SSHGD level courses. Secondary teachers were asked to classify the content of each item under one of a number of given headings. The list below is a simplified version of the list given the teachers, with some categories combined. Essentially, each teacher was asked whether the item content was:

- A. Old knowledge that students should have on entry to the course.
- B. New knowledge that all students are expected to learn in the course.
- C. New knowledge that some students are expected to learn in the course.
- D. New knowledge that no student is expected to learn.

University teachers were asked to make a similar assessment. In their case, the possible classifications were:

- A. Old knowledge that students should have on entry to the course.
- B. New knowledge that all students are expected to learn in the course.
- C. Other.

The responses did not fit a simple pattern. There were some items given the same classification by a large majority of teachers; there were others on which teachers were almost evenly spread over several classifications; and there were intermediate cases. Thus it was not possible to clearly identify many items as "content of the SSHGD course", for example, or "expected on university entrance". Responses to each item had to be considered individually.

When we move away from the test into the Project III area of program content, the problem is a different one. Secondary teachers and university teachers were given a list of topics which might be covered in an SSHGD level or first year university level physics course; this list is reproduced in Table 6.1. They were asked to state for each of these topics the average entry and

exit level of students on the following scale (these brief headings are defined in Table 6.2):

- 0 No knowledge
- 1 Descriptive knowledge
- 2 Introductory knowledge
- 3 Intermediate knowledge
- 4 Advanced knowledge

The project staff decided to work with the mean of teacher responses in this case. (The figures quoted in this report were compiled before all data were received; final figures were not available for Interproject Analysis.)

A further question to be considered in evaluating the usefulness and reliability of the data from both studies is that of the goodness of fit among the test used, the program laid out in the Ministry of Education guidelines (Circular G-S.17C, 1967), and the list of topics presented in the Course Description Questionnaire used in Project III. All the content of the standard physics program is included in the Project III list of topics. All the content of the test, with a very few minor exceptions, is included in the standard physics program, and is for the most part given major emphasis in that program. The Project III list of topics is considerably larger than the SSHGD program; this is to be expected, since it was set up to include not only topics that might be covered at this level, but also topics that might be covered at the first year university level. The standard program is of necessity more extensive than the test, first because it is impossible to test in 90 minutes all topics of that program, and second because the test was originally designed to be administered at a time of year when students would not yet have covered the portion of the course dealing with electricity and magnetism; these topics were

deliberately excluded. As well, the test concentrates most heavily on those parts of the course given major emphasis. Given these unavoidable restrictions, the test provides a reasonable coverage of all areas of the course except electricity and magnetism.

2. AN OVERVIEW OF THE DATA

Table 6.3 summarizes the most essential pieces of data resulting from Project II and Project III work. For each topic for which there were related test items, Table 6.3 provides the following:

- (a) the mean reported secondary exit level (on the 0 to 4 scale described above);
- (b) the mean reported university entrance level (on the same scale);
- (c) for each item related to the topic:
 - (i) the percentage of secondary respondents classifying the item as each of A, B, C and D (as described above);
 - (ii) the percentage of university respondents classifying the item as each of A, B and C;
 - (iii) the percentage of students correctly answering the item.

Some remarks must be made about the structure of this table before an attempt is made to do any interpretation of its contents. First, there are a number of items which occur in the table more than once. These items deal with more than one of the listed topics, and are tabulated under all topics with which they are concerned. A certain amount of caution is needed in

interpreting the student results on these items, since poor performance may be due to poor knowledge of any one of these topics.

Second, the only data considered here are those from the part of the study conducted in Anglophone schools, because the low number of respondents, particularly at the postsecondary level, in the Francophone part of the study resulted in data of insufficient precision to be useful. Information on test results in the Francophone schools, and on classification of the items by their teachers, may be found in the Project II report, but will not be further considered here.

With these cautions in mind, it is possible to derive some general impressions from Table 6.3 before moving on to deal with specific topics.

It will first be noted that there is a consistent and sizeable discrepancy (about two levels) between the level at which secondary teachers claim to send students out in June and the level at which university teachers claim to receive them in September. Over and over, secondary teachers reported an exit level of between 2 and 3, and sometimes even over 3 (introductory to intermediate knowledge and beyond). And over and over university teachers said the entrance level was between 0 and 1.5--not even introductory knowledge.

The obvious conclusion is that the two sets of teachers interpreted the scale in quite different ways, and that in rough terms what is considered intermediate knowledge by secondary teachers is considered introductory knowledge at the university level. The question then is, given this apparent difference of interpretation of the scale, what conclusions it is possible to draw from these figures.

It may be of some interest to see whether the two sets of teachers agreed on the relative amount of knowledge students have on different topics--that is, on which topics are best and least

understood. If for each set of teachers we rank-order the topics according to how they have been assessed on the scale, we find a large amount of agreement. Of the 38 topics, 20 appear in the same quartile in both rankings, and a further 12 are in adjacent quartiles. While this is certainly not unanimity, it does indicate highly similar perceptions of which topics are better known and which worse known.

A second general impression of the table is that there is a preponderance of items whose content is taught by the majority of teachers at both levels. This could indicate either of two situations--a duplication, so that students who already know certain material are being retaught it, or a gap, in which students have been taught the material but have not learned it, thus necessitating a second teaching. A glance at the actual difficulty levels of items fitting this description suggests that both situations occur, but that the second is more frequent.

The overall picture, then, is somewhat confused. Secondary teachers and university teachers seemed to be more or less in agreement about which topics students are most and least competent in, but differed greatly in their descriptions of the actual levels of competence. University teachers retaught a great deal of material already covered in secondary school; on the basis of the test results they seemed to be justified in this about half the time. Student knowledge, as manifested in the test results, appeared erratic--knowledge of some topics was good, of others poor, of still others variable.

3. INDIVIDUAL TOPICS

3.1 General Topic 1: Measurement

Topic 1.1, "Mass, length and time (fundamental quantities)", should, according to Ministry of Education guidelines, receive moderate emphasis at the SSHGD level. University teachers

considered this one of the relatively better handled topics, although secondary teachers placed it more than halfway down the list. The one test item on this topic deals with the stroboscopic effect. The content was assumed or taught by 77% of secondary teachers, and was assumed to be old knowledge by 75% of university teachers. However, the test results were poor, bordering on very poor.

Topic 1.5, "Development of relations relating physical quantities by dimensional analysis", warrants major emphasis according to the guidelines. Teachers at both levels considered this one of the most poorly handled topics. There are five related test items. The content of each of these items was covered by at least 80% of the secondary teachers, and considered old knowledge by from 60% to 100% of the university teachers. Student results were very poor on one item, which involved inverse proportion as well as presenting more than average opportunity for mathematical error. Performance was adequate on three items. Probable difficulties in these items included the use of ratio and proportion, the requirement that the student be familiar with units such as the newton, and the test's failure to supply formulas for the volume of a sphere and for centripetal force. In the fifth item, a relation was supplied and students had to apply it; performance on this item was good.

Of the six items dealing with measurement, results were good on one, adequate on three, poor on one, and very poor on one. Teacher consensus was that all the content of these items should have been learned by the end of the SSHGD course.

3.2 General Topic 2: Functions

All tested topics under this heading should receive major emphasis, according to the guidelines.

University teachers considered Topic 2.1, "Given a table of experimental data, plotting and labelling a graph of the data", relatively well handled, with secondary teachers rating it slightly lower. The one relevant test item required the student to know the relationship between the separation of nodal lines and wavelength, and to recognize the graph of a linear relation. The content was covered or assumed by 77% of secondary teachers. Student results were poor, but this is probably of little consequence for further work in physics, since only one-quarter of university teachers assumed student knowledge of this topic, and nearly half neither assumed it nor taught it in their courses.

Topic 2.2, "Given a table of experimental data, writing the equation of a linear relation", was also felt by university teachers to be relatively well handled, although secondary teachers placed it in the lower half of their rank-ordering. The relevant test item was considered review work at both levels; student performance was very good.

Topic 2.3, "Given a table of experimental data, replotting a non-linear relation to obtain a straight line and writing the relation", was felt by both sets of teachers to be poorly handled. The content of the one related test item was assumed or taught at the lower level, and assumed by 70% of teachers at the higher level, but student performance was just short of adequate.

Even though the Ministry of Education considers that these should be areas of major emphasis, the only item of the three dealing with functions on which performance was adequate was one reviewing pre-SSHGD level work.

3.3 General Topic 3: Motion (Kinematics)

All four of the topics in this section which are represented by test items are topics for major emphasis according to the guidelines.

Topic 3.1, "Motion with constant acceleration in a straight line", is relatively well handled by students according to teachers at both levels. The related test item requires students to deal with acceleration in terms of displacement, and to know Newton's Second Law. Performance was poor. The content was taught or assumed by almost all secondary teachers, and assumed by 70% of university teachers.

Topic 3.2, "Straight line kinematics with uniform acceleration", was also judged by both groups of teachers to be relatively well handled. There are eight test items related to this topic. The content of all was assumed or taught by at least 87% of secondary teachers, and in all but one case by at least 97% of these teachers. A minimum of 70% of university teachers considered each of these items as old material. Performance was good on two items, adequate on two more, poor on three, and very poor on one. The item with very poor results asked the student to deal with acceleration in terms of displacement, and to do so with variables rather than with given numbers. Of the poorly-done items, one required a knowledge of Newton's Second Law and of the result of removing a force on a frictionless surface. A second required the knowledge that under uniform acceleration, the average velocity over an interval is equal to the instantaneous velocity at the mid-point of the interval. The third required the student to distinguish between average velocity and average speed; to identify the type of motion from a displacement-time graph; and to derive velocity-time and acceleration-time graphs from a distance-time graph.

For Topic 3.3, "Straight line kinematics with non-uniform acceleration (e.g., circular motion)", both sets of teachers reported performance on the lower half of the scale. This is supported by the poor student performance on the item, whose content was assumed or taught by 87% of secondary teachers. Of university teachers, 40% assumed the knowledge, and a further 50% retaught it. The item dealt with a displacement-acceleration graph for a spring.

Secondary teachers felt students were relatively able to deal with Topic 3.5, "Two-dimensional kinematics with non-uniform acceleration (e.g., circular motion)"; university teachers ranked it slightly lower. Students performed much better on this topic than on the previous one, doing very well on two items, adequately on a further four, and poorly on two. One of the items with poor results had content taught or assumed by only 57% of secondary teachers; it required students to know the relation between centripetal acceleration and the radius of the circle. The content of the other "poor" item was dealt with by 86% of secondary teachers, although the item deals with a curved path rather than with the more familiar circular path. The item demands the use of vectors. One item where results were close to inadequate deals with an elliptical path. The three items with which students had the most success were taught or assumed by 92% or more of secondary teachers, and considered old knowledge by at least three-quarters of university teachers. The items on which performance was adequate had been taught or assumed by from 60% to 80% of secondary teachers. One was assumed by 70% of university teachers; for the other two, there was a spread of these teachers over all three classifications (old knowledge, course content, other).

Over the whole field of kinematics, results were mixed. Students performed very well on two items, well on two, adequately on six, poorly on seven, and very poorly on one. Most of the item content was considered old knowledge at the university level, although for three items a sizeable proportion of teachers treated the content as new (all three are items where results were poor).

3.4 General Topic 4: Newton's Laws of Motion--Dynamics of a Particle

On Topic 4.2, "Newton's Second Law", both secondary and university teachers reported student knowledge as relatively high. The guidelines recommend major emphasis on the topic. The content

of all six related items was covered by almost all secondary teachers (for only one item does the percentage fall below 97). The same pattern held at the university level--the content was considered old knowledge by 70% or more of teachers, except in one case where 40% assumed it and 55% taught it. Performance was variable--good on two items, adequate on one, poor on three. Content of the poorly-done items included the factors affecting momentum, the kinematics of constant acceleration, the effect of removing a force on a frictionless surface and the relationship between displacement and time under constant acceleration.

Topic 4.6, "Frames of reference", was placed by both sets of teachers near the bottom of the competency scale. The item content was widely taught at the SSHGD level, and widely assumed in universities. Student performance was adequate. (This topic is not mentioned in the guidelines.)

A mixed picture emerges here as well. Of seven items dealing with Newton's Laws, students performed well on two, adequately on two, and poorly on three.

3.5 General Topic 7: Gravity--Near the Earth's Surface

For the two topics in this area which are represented by test items, both secondary teachers and university teachers reported relatively high levels of knowledge.

Topic 7.2, "Weight and acceleration due to gravity", merits major emphasis according to the guidelines. It was represented by two test items; 87% and 97% of secondary teachers respectively assumed or taught their content. Most university teachers considered it old knowledge. Performance on one of the items was adequate, on the other poor. This latter item requires students to handle acceleration in terms of displacement, using variables rather than given numbers.

Topic 7.3, "The dynamics of projectile motion (no air resistance)", is recommended in the guidelines for moderate emphasis. Both relevant test items were widely covered by the end of the SSHGD course, and assumed by 70% or more of university teachers to be old knowledge. Performance on both items, which dealt with the basic principles of projectile motion, was poor.

There appears to be a problem in this area. According to their secondary teachers, students should be able to handle all these items. University teachers regarded the item content as previously acquired knowledge. Yet performance was adequate on only one of the four, poor on two, and very poor on one.

3.6 General Topic 8: Universal Gravitation

The guidelines suggest minor emphasis for Topic 8.3, "Universal Law of Gravitation". Both secondary and university teachers placed the topic near the mid-point of the rank-ordering of topics in terms of level of knowledge. The specific content of the one test item was almost universally covered in secondary school and widely assumed at university. Results on the item, which requires students to remember and deal with the above-mentioned law using a variable rather than a given number, were adequate.

Topic 8.4, "Circular orbits", should receive major emphasis according to the guidelines. Both groups of teachers placed student knowledge of this topic at or near the bottom in comparison with other topics. However, results on the one test item requiring students to deal with a circular orbit were adequate; the specific content of the item was widely taught at the secondary level and assumed by 60% of the university teachers (a further 35% taught it as new knowledge).

Thus both items on universal gravitation were dealt with adequately, even though one falls under a topic to be given minor emphasis.

3.7 General Topic 9: Momentum

Of the five topics under this heading which are represented by test items, all are recommended in the guidelines for major emphasis.

Three test items deal with Topic 9.1, "Impulse and momentum". Secondary teachers felt that this was a topic in which students were relatively highly competent; university teachers placed them slightly lower. One of the items was assumed or covered by virtually all secondary teachers, and considered old knowledge by two-thirds of university teachers; another quarter taught the material. Performance was good. On the other two items, which were less widely covered at the secondary level, performance was poor. One was assumed or taught to all students by 87% of secondary teachers, although all these teachers covered it with at least some of their students. Over half of the university teachers treated it as new knowledge, although 40% assumed it. The other was covered with all students or assumed by only 63% of teachers at the secondary level, though most covered it with some students. Only 20% of university teachers treated this as old knowledge; more than half taught the material, and 25% neither assumed nor taught it. Both poorly done items required the student to remember the equation for momentum. One required the student to find the area under a force-time graph with a non-constant force.

Teachers at both levels also felt students should be able to deal with Topic 9.2, "Conservation of linear momentum", at a fairly high level of competence. Results were adequate on one item, very poor on the second. The content of both items was covered at the earlier level by more than 80% of teachers, but taught as new material by a sizeable percentage of university teachers. The "very poor" item involved a speed-time graph.

University teacher assessments of student knowledge on Topic 9.3, "Elastic collisions (in one dimension)" placed it below the midpoint of the rank ordering, although secondary teachers placed

it near the top. The content of both test items was widely taught at the SSHGD level, but taught again by 40% to 50% of university teachers. Student performance was poor on one item, very poor on the other, which involved a speed-time graph.

Both groups of teachers felt students were relatively poorly able to deal with Topic 9.4, "Inelastic collisions (in one dimension)". The content of the one test item was widely taught at the lower level, but retaught by one-third of the university teachers. Results were adequate.

Topic 9.5, "Collisions of bodies in two dimensions", is one on which students have very little knowledge according to university teachers; secondary teachers gave a much higher estimate. Students did well on the one related item. The material was broadly taught in SSHGD courses, retaught by half the university teachers.

Eight items altogether deal with momentum (one deals with two different topics under this heading). Performance was good on two, adequate on one, poor on three, and very poor on two (dealing with elastic collisions and conservation of momentum).

3.8 General Topic 10: Work, Energy and Power

Again the items under this heading cover an area which the Ministry judges to warrant major emphasis. Secondary teachers ranked these topics in the top half of their list of relative student competencies. All test items on these topics have content which was covered at the secondary level with all students by at least 74% of teachers. In every case, the material was retaught by between 40% and 55% of university teachers.

The one item representing Topic 10.2, "Work done by a constant force", produced poor results. University teachers felt students were of less than average competence in this topic. The

item required students to work with an acceleration-distance graph for a spring.

Topic 10.5, "Kinetic energy and its relation to work", was considered by university teachers a topic where student knowledge was above the norm. Of three items on this topic, one produced poor results and two very poor results. Again, students were required to work with a speed-time graph, and with momentum without being supplied the formula.

There was only one test item on Topic 10.6, "Potential energy (constant force)"; results were adequate. University teachers placed student competence in the topic high.

Topic 10.7, "Potential energy (non-constant force)", was rated an area of very low competence by university teachers. Two items represented the topic. The results of one were poor, of the other very poor. The more poorly answered item dealt with potential energy in a spring; it did not provide the appropriate formula, and used variables rather than numbers. The other item involved universal gravitation.

University teachers felt students were relatively reasonably competent in dealing with Topic 10.8, "Conservation of mechanical energy (transformation between kinetic and potential energy)". Results on one of the three related items were good (this item included a "total energy" graph). On the other two, performance was poor. In one, variables rather than numbers were used; the other involved universal gravitation.

A total of nine items dealt with work, energy and power (again, one item dealt with two topics). Performance was good on one, adequate on one, poor on four, and very poor on three.

3.9 General Topic 11: Vibrations and Waves

Both university and secondary teachers assessed students' knowledge of Topic 11.12, "Relation between frequency, wavelength and velocity", as above the average. Major emphasis is suggested in the guidelines. Five test items deal with this topic. All were assumed or covered with all students by 90% or more of secondary teachers; all were retaught by a substantial number of university teachers, although they appeared to be irrelevant to the courses of a quarter or more of university teachers. One item was considered old knowledge at entry to the SSHGD course by 74% of secondary teachers, but taught as new material by one-third of those university teachers who classified the item as A or B. It required students to change units during their calculations; performance was poor. Results on the other four items were very good (1 item), adequate (1 item), and poor (2 items). One of the poor items also dealt with Snell's Law, on which results were generally poor.

Only one item represents Topic 11.13, "Waves propagated in one dimension". The guidelines suggest only minor emphasis on this topic, and both secondary and university teachers placed it below the midpoint on their scales of relative knowledge. The item content was widely covered at the secondary level, but retaught by two-fifths of the university teachers marking the item other than C. Student performance was adequate.

The two groups of teachers agreed that students have relatively little knowledge about Topic 11.16, "Doppler effect", which is not included in the guidelines. The content of the one test item on the topic was expected or taught by 71% of secondary teachers, retaught by nearly half of the university teachers to whose courses the topic was relevant. Student performance was poor. (It should be noted that it would be possible for an able student to answer this item even without specific instruction on the Doppler effect.)

In spite of the Ministry's judgement that major emphasis should be given to Topic 11.20, "Energy radiated by a point source (inverse square law)", teachers at both levels judged this to be one of the topics about which students knew least. The content of the one related item was covered by just over three-quarters of the secondary teachers, and taught as new material by a large number of those university teachers who dealt with the topic. Performance was very poor. One possible source of difficulty in the item is that the point source is a radioactive source rather than the more familiar source of light, or gravity; students may not have recognized the applicability of the Inverse Square Law.

The content of all eight items on vibrations and waves was covered at the secondary level; a high proportion of university teachers to whose courses the content was relevant retaught the material. Student test results were very good on one item, adequate on two, poor on four, and very poor on one.

3.10 General Topic 12: How Light Behaves

Topic 12.5, "Reflection of waves at a plane boundary", requires only qualitative treatment according to the guidelines; teacher assessment of student knowledge was relatively low at both levels. However, the content of the related test item was covered or assumed by just under 90% of secondary teachers, and considered old knowledge by nearly all physics teachers dealing with the subject. Student results on the item were poor; the item deals with reflecting barriers in a ripple tank which are set at an angle.

Moderate emphasis is suggested by the Ministry for Topic 12.9, "Concept of refractive index (Snell's Law)". University teachers felt this was one of the better-understood topics, although secondary teachers disagreed. 87% or more of secondary teachers in each case said that the content of each of the three related items should have been covered by the end of the SSHGD

course; many of the university teachers dealing with the topic treated it as new material. Student performance was poor on all three items, one of which also involved Topic 12.14 (see below).

Topic 12.14, "Refraction at a spherical interface", should receive only minor emphasis according to the guidelines; teachers at both levels rated student knowledge very low. As for the preceding topic, the content of the item on this topic was widely taught at the secondary level, but taught as new by a substantial number of university teachers. Performance was poor.

There were a total of four items involving General Topic 12; performance on these items was uniformly poor, in spite of the fact that the great proportion of secondary teachers had taught the material which they covered.

3.11 General Topic 13: Interference and Diffraction

The topics under this heading were generally rated comparatively low at both levels in terms of student knowledge.

One item deals with Topic 13.2, "Reflection and transmission of pulses and waves at a boundary", a topic the guidelines suggest should receive minor emphasis. The content of the related test item was covered by most secondary teachers, and assumed as old knowledge by most university teachers dealing with the subject. Performance was poor.

The guidelines suggest a qualitative treatment only for Topic 13.3, "Standing waves on a string, ends fixed". The related test item has content taught by most secondary teachers, and assumed by two-thirds of the university teachers treating the subject. Students did very well on this item.

Topic 13.6, "Interference of periodic waves, two point sources in a two-dimensional medium", warrants major emphasis according to the guidelines. There are three related test items.

Coverage with all students of item content varied at the secondary level between 62% and 90%, but in each case 90% or more of secondary teachers covered the material with at least some of their students. About one-half of the university teachers interested in the topic treated the item content as new knowledge. Test results were poor on two items, adequate on one (curiously, the one with the least widespread coverage in secondary school).

Major emphasis is also suggested in the guidelines for Topic 13.7, "Interference effects produced by a double slit (Young's double slit)". The content of the related item was widely taught at the secondary level, retaught at the university level by a substantial number of those teachers dealing with the topic. Student performance on this item was good.

Topic 13.10, "Fraunhofer diffraction by a single slit", should have, according to the guidelines, only minor emphasis. The one item on the topic has content somewhat less widely taught at the secondary level than was the case with most of the test items; 68% of teachers covered the material with all students, though almost 90% introduced at least some students to it. Well over half the university teachers dealing with the topic taught this item content as new material. Student performance was poor.

One item deals with Topic 13.19, "Interference effects in wedge-shaped thin films". Minor emphasis on this topic is suggested in the guidelines. The specific content of the item is not included in the guidelines, which deal with the topic only with respect to reflected light; the item deals with transmitted light. Coverage of this material at both levels was similar to that of the item related to the previous topic--two-thirds of secondary teachers covered it with all students, and almost 90% with some students, but almost two-thirds of interested university teachers retaught the material as new.

Performance varied widely on the eight items dealing with interference and diffraction. It was very good on one item, good on one, adequate on one, poor on four, and very poor on one.

4. GAPS

It is evident from the preceding section of this chapter that there exist portions of the secondary and postsecondary programs in physics in which students are expected to know at the higher level material they do not in fact know. It is the intention of this section to examine more closely those test items where this situation exists.

4.1 Gaps Related to Program

One reason for a gap is that students are not taught material at the lower level that teachers at the higher level expect on entry to their courses. There is evidence that this type of gap does occur in the situation under discussion.

Table 6.4 lists all test items whose content was covered with all students by 80% or fewer of secondary teachers, along with pertinent data about those items. This lack of coverage is, of course, only a problem if university teachers expect students to have learned the material. We have assumed in the preceding section that a university teacher who classified an item as C was very probably indicating that the content of the item had little or no relation to the content of his/her course. So there are two figures to be considered: the percentage of all university teachers expecting previous knowledge of the item content, and the percentage of university teachers dealing with the topic who expect this knowledge.

There are four items on this list whose content was expected on entry by 70% or more of all university teachers. A further six were expected by between 40% and 50% of these teachers. There can be little doubt that a gap exists with at least these ten items. If we examine the other relevant percentage--the percentage of teachers dealing with the topic who expect previous knowledge of the item content--we find that in three further cases this percentage exceeds 40. A perhaps less critical gap is indicated here as well.

Student performance on these items indicates that the problem is genuine. Of the ten items where the gap appears most serious, only three were answered adequately; the success rates were only just enough to put the items into the adequately answered class--31%, 32% and 30% respectively. Of the three somewhat less critical items, the students reached 35% success on one. The others were answered poorly and very poorly.

It might be suggested that the reason for less widespread coverage than is desirable, and/or for the poor test results, lies in the guidelines. However, with two exceptions these problem items deal with topics for which the guidelines suggest major emphasis. In these two cases (item 32 and 44), it is perhaps unreasonable for university expectations to be so high. But we are left with 11 items on topics recommended for major coverage, and yet with content not covered by 20% or more of SSHGD level teachers.

In two cases, the problem can plausibly be pushed one stop farther back. For items 3 and 33, one-third or more of SSHGD teachers assumed that the students had already learned the material before entering their courses. To judge by the test results, they had not, and the situation was not remedied during their final secondary year.

The two main problem areas appear to be Functions, where two of the three test items fall into this category, and Motion (Kinematics), where four such problems arise. Other general

topics involved are Measurement, Work, Energy and Power, Vibrations and Waves, and Interference and Diffraction.

4.2 Gaps Apparently Not Related to Program

There are a large number of test items the content of which was covered or considered old knowledge by more than 80% of secondary teachers, but on which student performance was very poor, poor, or just adequate. Table 6.5 lists all such test items which were answered correctly by fewer than 40% of the students. In all cases, a substantial percentage of university teachers considered the item content as old knowledge.

These items fall into two quite distinct categories. For 15 of the items, between one-third and three-quarters of SSHGD level teachers considered the item content already to be old knowledge on entry to their courses. For the remaining 12, considerably fewer SSHGD level teachers treated the content as previous knowledge (from 7% to 23%).

We will deal first with the 15 items treated as old knowledge at the SSHGD level. Four of these items deal with topics handled at a level of emphasis other than major (items 12, 13, 41 and 52), and it is fair to state that universities should have less high expectations that students know this content. Item 10 deals with three topics, one of which ("Frames of reference") is not included in the guidelines, although the other two receive major emphasis, and the "Frames of reference" component of the item does not appear difficult to handle without specific instruction. However, it is again probably unreasonable that university teachers should expect good knowledge of this item content.

The remaining 10 items all deal only with topics given major emphasis in the secondary course. In each case, at least one-third of the teachers of SSHGD level physics judged that students already had the knowledge necessary to do the item, and

yet fewer than 40% of the students could do so, and the success rate dropped as low as 11%.

As in the case of two similar items discussed in Section 4.1, there seems a clear indication that a serious problem exists. SSHGD level teachers of physics are assuming that students know considerably more on entry to their courses than they in fact do know, and deficiencies are often not corrected at that level. University teachers continue to assume that this material is known, and in large numbers do not teach it at their level either.

Considering only the 10 items on topics given major emphasis, we find that the most serious problems arise in the areas of Motion (Kinematics), where five items evince this sort of gap; Vibrations and Waves, where four of eight items are of this sort; and Newton's Laws of Motion--Dynamics of a Particle, where three of seven items show the same problem. Other areas involved are Measurement, and Gravity-- Near the Earth's Surface.

If we examine the 12 items whose content was widely taught for the first time at the SSHGD level, we discover that there are four of the items which deal with topics recommended for less than major emphasis at the level. As before, it seems reasonable to state that university teachers should not expect too much of students on these topics. The student success rate on these items ranged between 7% and 32%.

The topics dealt with by the majority of these items were in the areas of Momentum (six of eight items) and Work, Energy and Power (four of nine items). Also involved were Motion (Kinematics), Newton's Laws of Motion--Dynamics of a Particle, and Interference and Diffraction.

4.3 Summary

Table 6.6 presents in summary form the evidence that there are a number of gaps at the secondary-postsecondary interface in physics. The items suggesting gaps span all general topics covered by the test, with various topics showing a predominance of different types of gaps.

Problems in Measurement (General Topic 1) appear to result largely from over-optimistic assessments of student knowledge at the SSHGD entry point by secondary teachers, and by a lack of realization at the university level that some topics in this area are not given major emphasis in secondary school.

In Functions (General Topic 2), the difficulty appears to result from the fact that not enough teachers have covered the material dealt with by the items with all of their students

Different kinds of gaps appear in Motion (Kinematics) (General Topic 3), although most problems seem to result either from assumptions at the SSHGD level that students know more than they do on entry, or from the omission of the material from the courses of a number of secondary teachers.

The main difficulty with Newton's Law of Motion--Dynamics of a Particle (General Topic 4) is apparently that SSHGD teachers again assume previous knowledge which does not exist.

Gravity--Near the Earth's Surface (General Topic 7) contains topics treated with a less-than-major degree of emphasis. University expectations in this field are probably too high. The same situation exists with Universal Gravitation (General Topic 8).

On the other hand, problems arising under the heading of Momentum (General Topic 9) seem to be more often the result of students not mastering material which is new to them at the SSHGD level. This is also true of Work, Energy and Power

(General Topic 10), though in this case some of the problems result from teachers omitting the topics from their courses.

In the case of Vibrations and Waves, (General Topic 10) responsibility for the gaps is about equally divided between SSHGD teachers who assume students have previous knowledge which in fact does not exist, and university teachers who expect proficiency in an area not given major emphasis at the earlier level.

Unrealistic university expectations for content not emphasized in secondary school seems to be the problem in the area of How Light Behaves (General Topic 12).

Gaps in the field of Interference and Diffraction (General Topic 13) appear to result from a variety of causes. Some work is not emphasized at the secondary level and should not be taken for granted in university; some is taught at the SSHGD level but not learned; and some is considered old knowledge at that earlier level, even though students have not mastered it previously.

5. DUPLICATIONS

There are only three items on the test on which the students performed well or very well, and which were also reported to be taught as new material at the university level. All other items on which performance was good were for the most part considered already learned material by university teachers, and were taught as new by at most 25% of those teachers.

The items involved are number 26 (58% success, retaught by 50% of university teachers), number 28 (54% success, retaught by 50% of university teachers), and number 43 (67% success, retaught by 35% of university teachers). They deal respectively with Topics 9.5, 10.8 and 11.12, all areas given major emphasis in the guidelines. The first was treated as new

knowledge at the SSHGD level by all teachers; the second was considered new at that level by about three-quarters of the teachers; the third was assumed to be previous knowledge by just over half the teachers.

Because of the small number of these items, and their spread over three different general topics, little can be concluded other than duplication of program content at the two levels does not seem to be a serious problem.

6. SUMMARY

Of the 21 general topics included in the list used in Project III, the standard SSHGD Physics course deals to at least a limited extent with 13, frequently omitting a number of individual topics from a section. The test used by Project II contains items related to 11 of these 13 general topics, omitting only Electricity and Magnetism and Atomic Structure. Its 60 items touch on 38 of the 60 individual topics included in the course which fall under these eleven general headings.

On historical evidence, a somewhat arbitrary scale was established for item results. If 60% or more of students answered an item correctly, results were judged to be "very good". A 50% success rate was "good"; 30% was "adequate", 15% was "poor"; and below 15% was "very poor". Using this scale, student performance was very good on five items, good on seven, adequate on 15, poor on 26, and very poor on seven.

These judgements of item results must, of course, be modified by a number of factors. A poor success rate on an item whose content has not been taught to many students means something quite different from a similar result on content taught to everyone. Student performance on an item whose content most university teachers expect students to know is more crucial than on one whose content is not required knowledge on admission to a

first year university course. The data were therefore examined with a number of such factors in mind.

It was found that the problem of duplication--of material already known by the student being taught to him at the university level as new--was very rare, to judge by the test results. In only three cases was item content widely taught as new at that level when student performance on the item was good or very good.

A substantial number of items, however, suggested the existence of gaps between student knowledge on admission to the university and teacher expectations of the knowledge of incoming students. There appeared to be four different kinds of gaps, scattered over all topics tested. In some cases university teachers expected students to have command of material which was given little emphasis at the lower level in compliance with the Ministry of Education guidelines for the course. In others, even though these guidelines recommended major emphasis for a particular topic, 20% or more of secondary teachers did not cover the content of the relevant items with all of their students. On some topics, SSHGD teachers assumed a much higher level of previous knowledge from their entering students than was the case, if one may judge by the fact that students were often unable to handle test items whose content was described by secondary teachers as "old knowledge". The largest group of problem items, however, dealt with material widely taught at the SSHGD level, but whose content had not been mastered by the students even though guidelines suggested that the topic receive major emphasis.

If the gaps adumbrated by this analysis are real, then it appears that responsibility for them must be divided. Teachers at the earlier levels of the secondary school should be ensuring that students leave their courses with a good command of basic physical principles and relations. Those at the upper level should first of all be checking that their incoming students have this basic knowledge, which appears lacking in many cases, and then

reteaching the material where necessary. All teachers should be covering with the appropriate amount of attention and time those topics which the guidelines state to be worthy of major emphasis. And university teachers should not nourish unrealistic dreams of student competence in other areas of the subject, which of necessity must be short-changed if sufficient time is to be spent on essentials.

TABLE 6.1
TOPIC TABLE FOR PHYSICS

<u>Topic</u>	<u>Number of Related test items</u> *
1. <u>MEASUREMENT</u>	
1. Mass, length and time (fundamental quantities)	1
2. Errors	0
3. Accuracy and precision (significant figures, "rounding off")	0
4. SI	0
5. Development of relations relating physical quantities by dimensional analysis	5
6. Dimensional consistency of equations relating physical quantities	0
2. <u>FUNCTIONS</u>	
Given a table of experimental data:	
1. Plotting and properly labelling a graph of the data	1
2. Writing the equation of a linear relation	1
3. Replotting a non-linear relation to obtain a straight line and writing the relation	1
4. Plotting a power law relation on log-log paper and writing the relation	0
5. Plotting an exponential relation on semi-log paper and writing the relation	0

TABLE 6.1 (continued)

<u>Topic</u>	<u>Number of Related test items</u> *
<u>3. MOTION (Kinematics)</u>	
1. Motion with constant acceleration in one, two, or three dimensions	1
2. Straight line kinematics with uniform acceleration	8
3. Straight line kinematics with non-uniform acceleration	1
4. Two-dimensional kinematics with uniform acceleration (e.g., projectile motion)	0
5. Two-dimensional kinematics with non-uniform acceleration (e.g., circular motion)	8
6. Three-dimensional kinematics with uniform acceleration	0
7. Three-dimensional kinematics with non-uniform acceleration	0
<u>4. NEWTON'S LAWS OF MOTION - DYNAMICS OF A PARTICLE</u>	
1. Newton's first law, mass and inertia	0
2. Newton's second law	6
3. Newton's third law	0
4. Resolution and summation of forces by scale drawing	0
5. Resolution and summation of forces by analytical methods	0
6. Frames of reference	1
7. Pseudo forces	0
8. Dynamics of circular motion	0

TABLE 6.1 (continued)

<u>Topic</u>	<u>Number of Related test items</u> *
<u>5. STATICS</u>	
1. Moments	0
2. Laws of equilibrium	0
3. Simple machines, the lever, the incline plane, pulleys	0
<u>6. NEWTON'S LAW OF MOTION - DYNAMICS OF A RIGID BODY</u>	
1. Translational motion	0
2. Rotational motion, torque, moment of inertia, angular acceleration	0
3. Combined translational and rotational motion	0
4. Friction	0
<u>7. GRAVITY - NEAR THE EARTH'S CORE</u>	
1. Distinction between gravitational and inertial mass; principle of equivalence	0
2. Weight and acceleration due to gravity	2
3. The dynamics of projectile motion (no air resistance)	2
4. The dynamics of projectile motion (air resistance)	0
5. Dependence of g on distance from centre of earth or on latitude	0
<u>8. UNIVERSAL GRAVITATION</u>	
1. Ptolemy, Copernicus, Kepler	0
2. Kepler's Laws	0
3. Universal law of gravitation	1
4. Circular orbits	1
5. General motion under a central force	0

TABLE 6.1 (continued)

<u>Topic</u>	<u>Number of Related test items</u> *
9. <u>MOMENTUM</u>	
1. Impulse and momentum	3
2. Conservation of linear momentum	2
3. Elastic collisions (in one dimension)	2
4. Inelastic collisions (in one dimension)	1
5. Collisions of bodies in two dimensions	1
6. Angular momentum	0
7. Conservation of angular momentum	0
8. Precession	0
10. <u>WORK, ENERGY AND POWER</u>	
1. Work done by a constant force	0
2. Work done by a non-constant force (e.g., $F = kx$)	1
3. Work done by a non-constant force (e.g., $F \propto 1/r^2$)	0
4. Work performed by compressing a gas	0
5. Kinetic energy and its relation to work	3
6. Potential energy (constant force)	1
7. Potential energy (non-constant force)	2
8. Conservation of mechanical energy (transformations between kinetic and potential energy)	3
9. Power	0
10. Efficiency of work	0

TABLE 6.1 (continued)

<u>Topic</u>	<u>Number of Related test items</u> *
11. <u>VIBRATIONS AND WAVES</u>	
1. Vibrations	0
2. Kinematic description of SHM	0
3. Motion under a harmonic force ($F = -kx$)	0
4. Solutions of the equation of motion for an oscillating system (Newton's second law)	0
5. Mass on a spring	0
6. Simple pendulum (in small angle approximation)	0
7. Rigid pendulum (in small angle approximation)	0
8. Torsional oscillations	0
9. Conservation of energy in undamped oscillations	0
10. Exponential decay of oscillations	0
11. Derivation of the differential wave equation for compressional waves in a gas or transverse waves on a string	0
12. Relation between frequency, wavelength and velocity	5
13. Waves propagated in one dimension	1
14. Waves propagated in two dimensions	0
15. Polarization	0
16. Doppler effect	1
17. Shock waves	0
18. Energy carried by a wave in one dimension	0

TABLE 6.1 (continued)

<u>Topic</u>	<u>Number of Related test items</u> *
19. Relative and absolute intensity of sound waves (decibels)	0
20. Energy radiated by a point source (inverse square law)	1
 12. <u>HOW LIGHT BEHAVES</u>	
1. Qualitative discussion of light sources	0
2. Rectilinear propagation of light waves	0
3. Concept of a wave front	0
4. Huygens' principle	0
5. Reflection of waves at a plane boundary	1
6. Reflection of waves at a spherical boundary	0
7. Convex and concave mirrors	
a) scale drawings	0
b) analytical treatment	0
8. Refraction of waves at a plane interface between two media	0
9. Concept of refractive index (Snell's Law)	3
10. Total internal reflection	0
11. Refraction by a prism	0
12. Prism spectrometer--minimum deviation	0
13. Dispersive power of a medium	0
14. Refraction at a spherical interface	1
15. The lensmaker's equation	0
16. The thin lens equation	0

TABLE 6.1 (continued)

<u>Topic</u>	<u>Number of Related test items</u> *
17. Formation of images by lenses	
a) scale drawings	0
b) analytical treatment	0
18. Power of a lens	0
19. Chromatic aberration	0
20. Monochromatic aberrations	0
21. The eye and the camera	0
22. Eye defects and corrective lenses	0
23. Simple and compound microscopes	0
24. The telescope	0
 <u>13. INTERFERENCE AND DIFFRACTION</u>	
1. Superposition of pulses and/or waves	0
2. Reflection and transmission of pulses and waves at a boundary	1
3. Standing waves on a string, ends fixed	1
4. Standing waves in a pipe, both ends open	0
5. Standing waves in a pipe, one end open, one end closed	0
6. Interference of periodic waves, two point sources in a two-dimensional medium	3
7. Interference effects produced by a double slit (Young's double slit)	1
8. Interference effects produced by a multiple slit	0
9. Fraunhofer diffraction by a straight edge	0
10. Fraunhofer diffraction by a single slit (single slit interference)	1

TABLE 6.1 (continued)

<u>Topic</u>	<u>Number of Related test items</u> *
11. Fraunhofer diffraction by a circular aperture	0
12. Fraunhofer diffraction by a grating	0
13. Rayleigh resolution criterion	0
14. Resolving power of a grating	0
15. Grating spectrometer	0
16. Fresnel diffraction	0
17. Interference effects in parallel thin films	0
18. Michelson interferometer	0
19. Interference effects in wedge-shaped thin films	1

14. ELECTRICITY AND MAGNETISM

1. Electrostatics	0
2. Electric force (Coulomb's Law)	0
3. Electric field	0
4. Electric potential energy	0
5. Electric potential difference--volt	0
6. Sources of emf	0
7. Millikan experiment	0
8. Motion of a charge--ampere	0
9. Ohm's Law--constant resistance	0
10. Resistance of a conductor	0
11. D.C. electric circuits	0
12. Kirchhoff's Laws	0

TABLE 6.1 (continued)

<u>Topic</u>	<u>Number of Related test items</u> *
13. Capacitance	0
14. Properties of dielectrics	0
15. Transients (RC circuits)	0
16. Electric power in DC circuits	0
17. Magnetism	0
18. Magnetic field	0
19. Force on moving charge in a magnetic field	0
20. Ratio of charge to mass of the electron	0
21. Magnetic field produced by a moving charge	0
22. Force on a current carrying conductor in a uniform magnetic field	0
23. Electric meters--galvanometers	0
24. Electric motors	0
25. Emf in a conductor moving in a uniform magnetic field--Lenz' Law	0
26. Inductance	0
27. AC generator	0
28. DC generator	0
29. AC circuits	0
30. Transformers	0
31. Hysteresis	0
32. Electromagnetic spectrum	0

TABLE 6.1 (continued)

<u>Topic</u>	<u>Number of Related test items</u> *
<u>15. ATOMIC STRUCTURE</u>	
1. Thomson model of atom	0
2. Rutherford scattering experiment	0
3. "Solar system" model of atom (Rutherford model)	0
4. Properties of electron	0
5. Photoelectric effect (photons)	0
6. Compton effect	0
7. Particle-wave duality of radiation	0
8. Particle-wave duality of matter (de Broglie)	0
9. Line spectra (historical evidence, Balmer series, etc.)	0
10. Discrete energy levels (Franck-Hertz experiment)	0
11. Energy levels of hydrogen (Bohr model)	0
<u>16. NUCLEAR PHYSICS</u>	
1. Radioactive decay	0
2. Alpha, beta and gamma radiation: properties and spectra	0
3. Detection of radiation	0
4. Structure of the nucleus	0
5. Properties of nucleons	0
6. Nuclear reactions--general nature	0
7. Nuclear fission	0
8. Nuclear fusion	0
9. Radiation hazards	0

TABLE 6.1 (continued)

<u>Topic</u>	<u>Number of Related test items</u> *
<u>17. TEMPERATURE AND HEAT</u>	
1. Temperature	
a) scales	0
b) methods of measurement	0
2. Thermal expansion	0
3. Heat	
a) kinetic theory	0
b) Maxwellian velocity distribution	0
c) conversion of mechanical to thermal energy	0
d) specific heat	0
e) calorimetry	0
f) gas laws	0
g) change of phase	0
h) vapour pressure and humidity	0
4. Heat transfer	
a) convection	0
b) conduction	0
c) radiation	0
5. Thermodynamics	
a) first law	
b) second law	
c) the Carnot cycle	0
<u>18. PROPERTIES OF SOLIDS OTHER THAN THERMAL</u>	
1. Crystallographic properties of simple solids	0
2. Elastic properties, Hooke's law, moduli and compliances	0
3. Electronic properties; band structure; conductors; semi-conductors and insulators	0
4. Electronic properties, the diode	0
5. Electronic properties, the transistor	0

TABLE 6.1 (continued)

<u>Topic</u>	<u>Number of Related test items</u> *
19. <u>FLUIDS AT REST AND IN MOTION</u>	
1. Density and specific gravity	0
2. Atmospheric pressure--the barometer	0
3. Hydrostatic pressure--Pascal's Law	0
4. Archimedes' principle--buoyancy	0
5. Surface tension and capillary action	0
6. Fluid, flow, continuity conditions	0
7. Streamline flow	0
8. Bernoulli's Principle	0
9. Turbulent flow	0
10. Viscosity	0
11. Poiseuille's Law	0
20. <u>SPECIAL THEORY OF RELATIVITY</u>	0
21. <u>PROPERTIES OF ELEMENTARY PARTICLES</u> (other than proton, neutron or electron)	0
22. <u>OTHER TOPICS</u> (please list and respond appropriately)	

*Some items are related to more than one topic.

TABLE 6.2
RATING SCALE FOR PHYSICS TOPICS

0 -No Knowledge

No knowledge of the topic is assumed on entrance or the topic is not covered in this course.

1 -Descriptive

Students understand the topic in a purely descriptive manner. They are capable of a verbal description, including appropriate definitions, and have some general idea of where the topic fits in the overall subject.

2 -Introductory

Students understand relevant concepts regarding the topic including any theoretical treatment required. NO vector algebra or calculus is used or assumed. Students are able to provide descriptive answers to SIMPLE questions, and if appropriate, are able to do exercises requiring SIMPLE algebraic manipulations and/or SIMPLE graphical methods.

3 -Intermediate

Students understand relevant concepts regarding the topic. They can use SIMPLE vector algebra and INTRODUCTORY calculus if appropriate. Students are able to provide descriptive answers to questions, and if appropriate, to do NUMERICAL EXERCISES involving one or more of the following techniques: algebraic manipulation, vectors, and graphical methods.

4 -Advanced

Students understand relevant concepts regarding the topic. They can use vector algebra and calculus if appropriate. Students are able to discuss this topic in descriptive terms, and if appropriate, are able to SOLVE PROBLEMS stated in descriptive terms with or without numerical data included, using one or more of the following techniques: algebraic manipulation, vectors, calculus and graphic analysis.

TABLE 6.3

PHYSICS: SUMMARY OF ESSENTIAL DATA

Topic No.	Reported secondary exit level*	Reported university entrance level*	Item No.	Secondary classification(%)**			University classification(%)**			% correct
				A	B	C	A	B	C	
1.1	2.5	1.2	32	11	66	8	15	10	15	15
1.5	2.3	0.3	2	64	27	3	5	0	0	52
			3	33	47	7	14	11	11	13
			23	4	85	11	0	35	5	32
			31	59	26	5	10	0	0	33
			51	34	63	1	1	25	5	34
2.1	2.9	1.1	19	8	69	18	4	30	45	25
2.2	2.8	1.4	1	77	19	1	3	0	0	70
2.3	2.6	0.3	33	38	38	17	8	15	5	29
3.1	3.1	1.5	49	37	60	3	0	20	10	21
3.2	3.2	1.6	4	49	49	3	0	10	5	25
			5	59	40	1	0	5	5	39
			6	58	41	1	0	5	5	29
			10	45	42	11	1	10	20	32
			21	33	67	0	0	15	5	22
			22	30	67	1	1	10	5	11
			34	38	60	1	0	15	5	50
			35	37	60	3	0	15	5	50

TABLE 6.3 (continued)

Topic No.	Reported secondary exit level*	Reported university entrance level*	Item No.	Secondary classification(%)**			University classification(%)**			% correct	
				A	B	C	A	B	C		
3.3	2.6	0.6	50	5	82	10	3	40	50	10	19
3.5	3.0	0.9	7	22	75	3	0	80	15	5	69
			8	12	84	4	0	75	20	5	67
			9	7	85	8	0	75	20	5	44
			24	6	71	14	10	45	25	30	31
			25	16	63	12	8	45	25	30	32
			37	11	75	10	4	50	45	5	23
			38	22	41	14	23	70	20	10	30
			39	1	56	19	23	50	35	15	17
4.2	3.1	1.4	21	33	67	0	0	80	15	5	22
			34	38	60	1	0	80	15	5	50
			35	37	60	3	0	80	15	5	50
			49	37	60	3	0	70	20	10	21
			51	34	63	1	1	70	25	5	34
			60	6	83	11	0	40	55	5	29
4.6	2.2	0.3	10	45	42	11	1	70	10	20	32
7.2	3.0	1.1	10	45	42	11	1	70	10	20	32
			22	30	67	1	1	85	10	5	11
7.3	3.1	1.0	53	23	73	4	0	75	20	5	23
			54	7	88	5	0	70	25	5	17

TABLE 6.3 (continued)

Topic No.	Reported secondary exit level*	Reported university entrance level*	Item No.	Secondary classification(%)**			University classification(%)**			% correct
				A	B	C	A	B	C	
8.3	2.8	0.7	52	37	58	3	75	20	5	32
8.4	0.4	0.2	23	4	85	11	60	35	5	32
9.1	3.0	0.9	20	10	85	5	65	25	10	52
			27	7	56	29	20	55	25	23
			60	6	83	11	40	55	5	29
9.2	3.1	0.8	36	7	86	7	55	35	10	32
			57	8	75	15	40	50	10	13
9.3	3.0	0.6	55	3	84	12	40	50	10	28
			56	7	81	10	45	40	15	7
9.4	2.8	0.7	36	7	86	7	55	35	10	32
9.5	2.9	0.3	26	0	88	11	40	50	10	58
10.2	2.9	0.4	50	5	82	10	40	50	10	19
10.5	2.9	0.8	56	7	81	10	45	40	10	7
			57	8	75	15	40	50	10	13
			60	6	83	11	40	55	5	29

TABLE 6.3 (continued)

Topic No.	Reported secondary exit level*	Reported university entrance level*	Item No.	Secondary classification(%)**				University classification(%)**			% correct
				A	B	C	D	A	B	C	
10.6	3.0	1.2	29	46	50	3	1	55	40	5	36
10.7	2.9	0.2	58	3	75	16	5	40	50	10	7
			59	4	70	14	12	25	55	20	16
10.8	2.9	0.8	28	23	68	8	0	45	50	5	54
			30	15	64	19	1	50	45	5	20
			59	4	70	14	12	25	55	20	16
11.12	2.9	1.1	11	51	40	5	4	55	20	25	32
			12	48	42	5	4	50	25	25	19
			14	40	48	5	7	45	20	35	21
			16	74	22	1	3	50	25	25	20
			43	52	44	1	3	40	35	25	67
11.13	2.7	0.7	13	58	32	4	7	45	30	25	34
11.16	1.5	0.4	44	18	54	17	11	40	35	25	25
11.20	1.4	0.2	42	7	70	16	7	35	25	40	14
12.5	2.4	0.5	45	16	73	8	3	50	10	40	23
12.9	2.8	1.1	12	48	42	5	4	50	25	25	19
			40	23	68	4	4	45	25	30	17
			41	47	40	10	4	35	30	35	21

TABLE 6.3 (continued)

Topic No.	Reported secondary exit level*	Reported university entrance level*	Item No.	Secondary classification(%)**			University classification(%)**			% correct
				A	B	C	A	B	C	
12.14	0.5	0.2	41	47	40	10	35	30	35	21
13.2	2.6	0.6	45	16	73	8	50	10	40	23
13.3	2.3	0.4	15	71	23	3	50	25	25	64
13.6	2.9	0.6	18	8	82	8	35	30	35	20
			19	8	69	18	25	30	45	25
			47	5	67	19	30	25	45	35
13.7	2.9	0.8	17	5	86	7	35	20	45	57
13.10	2.4	0.1	48	4	64	21	20	35	45	16
13.19	2.4	0.1	46	5	62	21	20	35	45	11

*Preliminary figures (see text, Section 1).
See Table 6.2 for description of levels.

**See text, Section 1, for explanation of classifications.

TABLE 6.4

TEST ITEMS WHOSE CONTENT IS COVERED WITH ALL STUDENTS
BY 80% OR FEWER OF SECONDARY TEACHERS

Item No.	Topic(s)	Secondary classification (%)**				University classification (%)**			Ministry emphasis	% Correct
		A	B	C	D	A	B	C		

I. Content expected on entry by 40 per cent or more of all university teachers.

3	1.5	33	47	7	14	79	11	11	major	13
24	3.5	6	71	14	10	45	25	30	major	31
25	3.5	16	63	12	8	45	25	30	major	32
30	10.8	15	64	19	1	50	45	5	major	20
32	1.1	11	66	8	15	75	10	15	moderate	15
33	2.3	38	38	17	8	80	15	5	major	29
38	3.5	22	41	14	23	70	20	10	major	30
39	3.5	1	56	19	23	50	35	15	major	17
44	11.16	18	54	17	11	40	35	25	*	25
58	10.7	3	75	16	5	40	50	10	major	7

TABLE 6.4 (continued)

Item No.	Topic(s)	Secondary classification (%)**				University classification (%)**			Ministry emphasis	% Correct
		A	B	C	D	A	B	C		
II. <u>Content expected on entry by 40 per cent or more of university teachers dealing with material.</u>										
19	2.1, 13.6	8	69	18	4	25	30	45	major	25
42	11.20	7	70	16	7	35	25	40	major	14
47	13.6	5	67	19	8	30	25	45	major	35
III. <u>Content expected on entry by fewer than 40 per cent of university teachers dealing with material.</u>										
27	9.1	7	56	29	8	20	55	25	major	23
46	13.19	5	62	21	12	20	35	45	minor	11
48	13.10	4	64	21	11	20	35	45	minor	16
59	10.7, 10.8	4	70	14	12	25	55	20	major	16

*Not included in guidelines.

**See text, Section 1, for explanation of classifications.

TABLE 6.5

TEST ITEMS ANSWERED CORRECTLY BY FEWER THAN 40% OF STUDENTS
BUT WHOSE CONTENT IS COVERED BY MORE THAN 80% OF SECONDARY TEACHERS

Item No.	Topic(s)	Secondary classification (%)***				University classification (%)***			Ministry emphasis*	% Correct
		A	B	C	D	A	B	C		
I. <u>Content expected on entry to SSHGD course by one-third or more of secondary teachers.</u>										
4	3.2	49	49	3	0	85	10	5	major	25
6	3.2	58	41	1	0	90	5	5	major	29
10	3.2, 4.6, 7.2	45	42	11	1	70	10	20	major**	32
11	11.12	51	40	5	4	55	20	25	major	32
12	11.12, 12.9	48	42	5	4	50	25	25	major- moderate	19
13	11.13	58	32	4	7	45	30	25	qualitative only	34
14	11.12	40	48	5	7	45	20	35	major	21
16	11.12	74	22	1	3	50	25	25	major	20
21	3.2, 4.2	33	67	1	1	80	15	5	major	22

TABLE 6.5 (continued)

Item No.	Topic(s)	Secondary classification (%)***				University classification (%)***			Ministry emphasis*	% Correct
		classification (%)***				classification (%)***				
		A	B	C	D	A	B	C		
22	3.2, 7.2	30	67	1	1	85	10	5	major	11
31	1.5	59	26	5	10	100	0	0	major	33
41	12.9, 12.14	47	40	10	4	35	30	35	moderate- qualitative only	21
49	3.1, 4.2	37	60	3	0	70	20	10	major	21
51	1.5, 4.2	34	63	1	1	70	25	5	major	34
52	8.3	37	58	3	3	75	20	5	minor	32

II. Content expected on entry to SSHGD course by fewer than one-third of
secondary teachers.

17	7.3	7	88	5	0	70	25	5	moderate	17
18	13.6	8	82	8	1	35	30	35	major	20
23	1.5, 7.3, 8.4	4	85	11	0	60	35	5	major- moderate	32
36	9.2, 9.4	7	86	7	0	55	35	10	major	32
37	3.5	11	75	10	4	50	45	4	major	23

TABLE 6.5 (continued)

Item No.	Topic(s)	Secondary classification (%)***				University classification (%)***			Ministry emphasis*	% Correct
		A	B	C	D	A	B	C		
40	12.9	23	68	4	4	45	25	30	moderate	27
45	12.5, 13.2	16	73	8	3	50	10	40	minor- qualitative only	23
50	3.3, 10.2	5	82	10	3	40	50	10	major	19
55	9.3	3	84	12	1	40	50	10	major	28
56	9.3, 10.5	7	81	10	3	45	40	15	major	7
57	9.2, 10.5	8	75	15	1	40	50	10	major	13
60	4.2,9.1,10.5	6	83	11	0	40	55	5	major	29

*Where suggested emphasis differs for the topics dealt with by the item, all relevant emphases are listed.

**Topic 4.6 is not included in the guidelines; the other two are given major emphasis.

***See text, Section 1, for explanation of classifications.

TABLE 6.6
DISTRIBUTION OF ITEMS SUGGESTING GAPS

General Topic	Total Items	Type of gap*			
		1	2	3	4
1	6	1	2	0	2
2	3	2	0	0	0
3	18	4	5	2	1
4	7	0	3	1	1
7	4	0	1	0	3
8	2	0	0	0	2
9	8	0	1	4	0
10	9	2	0	4	0
11	8	1	3	0	3
12	4	0	0	0	4
13	8	2	0	1	1

*Types of gap:

1--major emphasis; expected at university; covered with students by 80% or fewer of secondary teachers

2--major emphasis; expected at university; taught at SSHGD level or assumed by more than 80% of secondary teachers; assumed on SSHGD entry by one-third or more of secondary teachers

3--major emphasis; expected at university; taught at SSHGD level or assumed at SSHGD entry by fewer than one-third of secondary teachers

4--other than major emphasis

CHAPTER SEVEN

FRENCH AS A SECOND LANGUAGE*

1. INTRODUCTION

In recent years, there have been substantial changes in the emphasis given to various skills in secondary school courses in French at the senior level. The Project III report for this study suggests that these changes include a reduction in the amount of attention paid to grammar, writing and translation (particularly the last of these), and a corresponding increase in the relative importance of the spoken language and of literature. The report also points out that there is a great deal of variability from school to school in course content and emphases, and that teachers of Grade Thirteen courses reported a similar variability in the levels of knowledge and capability of their incoming students. University teachers indicated that this variability is still present on university entry. Overall, teachers at both these levels reported that they were somewhat dissatisfied with the skills of their incoming students.

The Course Description Questionnaires used by Project III to elicit from secondary and university teachers a description of the content of their courses and the relative importance given to various skills included a list of 17 "topics"--perhaps more accurately, in the case of French, described as "skills". This list is given in Table 7.1. It was not possible for the Project

*(This chapter was written with the assistance of Mary Hainsworth, a part-time researcher.)

II staff to test all these topics, and in fact no attempt was made to do so. Partly because of the variability mentioned above in the content of Grade Thirteen courses, attempts to find or develop a test closely tied to this content would have served little purpose. The tests used in the study are more properly considered as measuring the students' degree of facility in using French in various ways than as measuring their achievement in a particular course.

This does not mean, however, that the tests are totally unrelated to the secondary school curriculum. Teachers of Grade Thirteen French (along with teachers of first year university courses in French) were asked to examine the tests in terms of the expectations they held of what students should know, both on entry to and on exit from their courses. Detailed data on their responses are reported in Appendix A3 of the Project II report; some of these responses are also included in this chapter at appropriate points. In general, teachers judged the tests to be appropriate in kind and in degree of difficulty for testing the particular skills on which they focussed, and (as is discussed later in this chapter) also agreed that developing these skills was a central goal of their courses. Some reservations were expressed or implied about particular parts of particular tests.

A battery of four tests was used in Project II, with each test addressing itself primarily to a different skill, although indirectly to others. The first was the Reading Test, which was most concerned with Topic 11, "Reading comprehension", though it dealt unavoidably with Topic 8, "Vocabulary--general", as well. The Listening Test dealt most directly with Topic 12, "Aural comprehension", and again necessarily involved Topic 8. The Writing Test measured competencies in Topics 1 to 5--knowledge of various grammatical points, ability to apply this knowledge in writing and speaking, and ability to express ideas clearly and correctly in written French. Inevitably, vocabulary was also tested. The Speaking Test measured the same grammatical competencies, the ability to apply them, and fluency in spoken French (Topic 13), along with vocabulary.

It is important to note that, although the testing done by Project II did not cover the entire list of skills examined by Project III, it did deal with those skills judged most important by teachers participating in Project III. There are several pieces of data corroborating this statement in the Project III findings. First, secondary teachers were given a list of 14 possible aims of a Grade 13 French course, and asked how much emphasis they gave these aims in their course. Knowledge of grammar was not specifically included in this list. Of the 14 aims listed, there were only six to which almost all teachers reported giving moderate to heavy emphasis (94% or more of the teachers questioned--the next highest number was 84%). Four of these six aims were: skill in reading; skill in writing; aural comprehension; and skill in speaking. (The others were respect for precision of expression--oral and/or written--and self-confidence in French.) Second, the same teachers were asked to state the proportion of class time they allocated to each of the 17 topics listed in Table 7.1. The topics receiving on the average 10% or more of class time were: Topics 1 to 4, considered as a unit (grammar)--22%; Topic 5 (expressing ideas clearly and correctly in written French)--13%; Topics 8 to 10 as a unit (vocabulary)--10%; Topic 11 (reading comprehension)--13%; Topic 12 (aural comprehension)--10%; and Topic 13 (fluency in spoken French)--13%.

The topics examined by these tests, then, as a group occupied on the average 68% of classroom time at the Grade Thirteen level, and were those judged by teachers at that level to warrant most emphasis. The same pattern held at the university level. The aims given moderate to heavy emphasis by most teachers at this level (85% or more of those questioned) were the same as those at the lower level, with the addition of only one--"appreciation of language as a medium of human thought". The topics given the largest proportion of class time were grammar (18%); fluency in spoken French (14%); ability to express ideas clearly and correctly in spoken French (14%); aural comprehension (13%); and vocabulary (8%). The total allotment of

class time to these topics is thus 67%, just under the secondary figure of 68%.

2. PROJECT III DATA

Project III was responsible for developing the list of topics to which we have been referring. We have also used data from this project to confirm that the areas examined by the Project II tests are in fact those considered most important by both secondary and university teachers of French.

One more piece of information is available to this report from the Project III work. For each of the 17 topics included in the list, both secondary and university teachers were asked to locate, on a 0 to 4 scale, the following:

- (a) the level at which, on the average, students entered their course
- (b) the level at which, on the average, students left their course.
- (c) the level at which they would prefer to have students enter their courses.
- (d) the level at which, given this preferred level, they could send students out of their course.

In this report, the only figures that will be examined are secondary teachers' assessments of their students' actual exit level, and the university teachers' assessments of their students' actual entrance level. The mean reported levels for each topic, along with the definition of the scale, are given in Table 7.2.

One is immediately struck in looking at this table by the differences in these levels. The university teachers were considering students coming from the secondary teachers' classes, and one would assume that in fact they were looking at the better students from those classes. Yet their assessment of student knowledge is in every case lower than the assessment made by the secondary teachers--by about one level, on the average. (This is a pattern that recurs in all subjects, rather than being peculiar to French.) The only logical conclusion seems to be that the two sets of teachers had a different understanding of the scale--that what a secondary teacher considered, for example, moderate competence, was considered by the university teacher to be minimal competence. These figures as they stand, then, tell us very little. However, if they are used to rank-order the topics--to tell us which topics each group of teachers considered best-handled and worst-handled by the students--a very high degree of consistency emerges. Of the six topics ranked highest by each group of teachers, five correspond. The same five topics are ranked lowest by each group. And in general, rankings correspond very closely. The largest discrepancies occur with topics 7 and 8. University teachers have a somewhat higher opinion of students' skill in translating from French to English than do secondary teachers, and the reverse is true in the case of general vocabulary.

3. PROJECT II DATA

The project II data on French as a second language are much less homogeneous in form than, for example, the data on mathematics and physics. A detailed discussion of these data will be left to the various appropriate spots later in the chapter. There are, however, some general points that may be made here concerning the relationship between the data and the topic list.

Topics 1 to 4 inclusive deal with grammar. Grammar is not tested in either the Reading Test or the Listening Test, except insofar as in isolated cases a knowledge of a particular grammatical point may be required if the student is to understand what he/she is reading or hearing. It is tested in the Writing Test and Speaking Test, but most often in a way which prevents the investigator from separating out these four topics. The structure of the testing is such that, although one can report on the students' level of performance with respect to grammar in general, it is very difficult or impossible to talk specifically about their competence with, for instance, verb tenses and moods. As well, the distinction between competence in grammar and competence in applying grammar in writing and speaking is one that in practice it is difficult to make, since this application of grammar to writing and speaking is the only context we have in which to measure the students' theoretical knowledge of grammatical points.

A more difficult problem arises with Topic 8 (Vocabulary--general). Vocabulary is tested throughout all four tests in the battery, but never alone. There seems no way to isolate this component of the test so that any conclusions can be made about the extent and range of students' vocabulary. The best that can be done is to point out that student performance was very good on many test items, and one must assume that this implies a knowledge of the vocabulary used in these items. No attempt will be made, therefore, to interpret any of these test results as indicating a particular level of knowledge of vocabulary. (This may not be such a serious omission as it might appear, since, to quote a comment made by one teacher in reaction to the tests, "Vocabulary is a rather chancy subject since there are no standard words to be learned.")

What will be examined, then, are the Project II data relating to grammar (Topics 1 to 4); ability to express ideas clearly and correctly in written French (Topic 5); reading comprehension (Topic 11); aural comprehension (Topic 12); and fluency in spoken French (Topic 13).

4. READING

4.1 Reading Test, Part A

The Reading Test consisted of two parts. In Part A, students were given a number of sentences from which a word or group of words had been omitted, and were asked to choose the correct word(s) from among four possible responses.

Teachers at both levels were asked to comment on each of these items individually, assessing on a scale supplied to them the elements of vocabulary and grammar needed to respond to the item. The possible responses for secondary teachers were:

A. Old knowledge that students should have on entry to the course.

A1. This knowledge is not reviewed in the course.

A2. This knowledge is reviewed in the course.

B. New knowledge that all students are expected to learn in the course.

C. New knowledge that some students are expected to learn in the course.

C1. Only 1% to 25% of students should learn this.

C2. Only 26% to 50% of students should learn this.

C3. Only 51% to 75% of students should learn this.

C4. More than 75% but not all students should learn this.

D. New knowledge that no student is expected to learn.

For university teachers, categories A and B remained the same; C and D were replaced by:

C. Other.

Table 7.3 gives, for each item in Part A of this test, the following information:

- (a) the percentage of students correctly answering the item;
- (b) the percentage of secondary teachers classifying the item as each of A, B, C and D;
- (c) the percentage of university teachers classifying the item as each of A, B and C.

These items can be divided, using the criterion of percentage of students correctly answering each item, into two groups. Nine of the items were correctly answered by 60% or more of the students. The remaining eight items had a success rate of under 50%. (No item fell into the 50% to 59% correct range.) It may be of interest to see whether there is a corresponding division of the teacher assessments at both the secondary and university levels.

We examine first the percentage of teachers at the secondary level describing these items as A or B--that is, knowledge that students should have no later than the end of Grade Thirteen. For the nine better-answered items, the mean percentage is 87; for the eight more poorly answered items, the mean percentage is 60. This is a substantial difference, and suggests that the secondary teachers have a fairly clear idea of what their students can and cannot do. The evidence is even stronger when it is pointed out that in only three cases of the nine where performance was high

does this percentage fall below 85, and that on the eight more difficult items in only two cases does it rise above 65.

If we examine in the same way the statements of university teachers about whether students should have the knowledge in these items on entry to their courses (the A ratings), we find the same sort of pattern. For the nine well-done items, the mean percentage of A ratings is 79; for the eight poorly-done ones, the mean percentage is 47. Although these figures are lower than the corresponding ones for secondary teachers, they differ by about the same amount from one another as do the secondary figures. On only three of the well-done items does the percentage fall below 75; on only two of the poorly-done ones does it exceed 50.

What is even more interesting is that if we examine the five anomalous items in each case (those apparently misjudged by university teachers), we find that four of the five items correspond. The three items on which performance was good, but which received fewer than 85% A or B ratings from secondary teachers, are the same three items that received fewer than 75% A ratings from university teachers. The two groups agreed on one item which, although poorly done, was given more than 65% A or B ratings by secondary teachers and more than 50% A ratings by university teachers. Each group included one more item in this category; these do not correspond, but in each case the item in question came very close to the borderline on the other list.

In other words, then, the secondary and university teachers generally agreed on what students should know at the end of Grade Thirteen and the beginning of university. However, of the 17 items in this part of the Reading Test, there were six which indicated by their results that both groups of teachers had made the same misjudgement about a particular skill or piece of knowledge. In three cases (Items 7, 11 and 13) students were performing better than their teachers expected, and in three cases (Items 6, 10 and 16) more poorly. It is of interest that two of the three items on which students "over-performed" dealt with

idiomatic constructions, and that two of the three on which they "under-performed" concerned reflexive or semi-reflexive verbs.

4.2 Reading Test, Part B

The second part of the Reading Test contained six short passages in French, each accompanied by from three to five questions about the passage in multiple-choice four-option form. Teachers at both levels were asked to comment on the difficulty of each of these passages for students at the level being tested. They were also asked to state the percentage of their students, both at entry to and on exit from their courses, who should be able to read a passage for literal understanding, to identify the main idea or purpose, and to draw inferences and see implications, and to specify the degree of emphasis they placed on each of these skills in their teaching.

(It is difficult to distinguish clearly which test items test each of the three skills mentioned above. The majority of the items, particularly on the earlier passages, test literal understanding; some, particularly on the later passages, ask for some inference, although not at a high level. No items test primarily the students' ability to grasp the main idea or purpose of a passage. No effort will be made here to compare test performance on those items with an element of inference and those which lack this element, other than to mention that the passage on which performance was worst--number 5--was also the one whose items demanded the greatest degree of inference.)

Table 7.4 reports the responses of both sets of teachers to the general questions, and Table 7.5 their assessment of the difficulty of the various passages. For each passage, Table 7.5 also includes the number of students correctly responding to each question related to the passage.

It can be seen from Table 7.4 that one-half of the Grade Thirteen teachers felt that all their students should, on leaving Grade Thirteen, be able to read a passage such as those in the test for literal understanding, and that 88% of these teachers had this expectation for over three-quarters of their students. The comparable figures for university teachers at entry to their courses are: 50% expected all students to have this skill, and 78% expected at least three-quarters of their students to have it. There is not a great discrepancy in these figures, although university expectations are slightly lower than secondary expectations.

When it came to identifying the main idea or purpose of such a passage, there was a much greater difference in expectations. Almost all secondary teachers felt that at least three-quarters of their outgoing students should be able to do this, but only 61% of university teachers had the same level of expectation of their incoming students.

The gap in expectations closed somewhat with respect to drawing inferences and seeing implications. A much smaller percentage of both groups expected students to have this skill at the interface--54% of secondary teachers expected this of three-quarters or more of their outgoing students, and 45% of university teachers expected it on entry to their courses from three-quarters or more of students.

This difference in expectation reveals itself as well in Table 7.5. Without exception, university teachers considered the test passages more difficult than did the secondary teachers, although their rankings of these passages according to difficulty corresponded closely. The assessments given in this table were arrived at by averaging teacher responses on each level, with 1 representing "too easy", 2 "somewhat easy", 3 "about right", 4 "somewhat difficult" and 5 "too difficult". Overall, two of the passages were assessed by both groups as in the "somewhat easy" range, two in the "about right" range, and two as approaching "somewhat difficult".

Again, there seems a reasonable correspondence between the expectations of the two groups of teachers, with the university teachers being slightly more pessimistic than the secondary teachers about student capabilities. These expectations are also relatively well borne out by student performance. Passages 1 and 2, rated "somewhat easy", produced mean success rates on the related items of 91% and 87% respectively. Passages 5 and 6, judged the most difficult by both groups, had mean success rates of 41% and 63% respectively. One of the "about right" passages had a mean success rate of 69%--approximately what would be expected in view of the other results. However, students performed much better comparatively on the other middle difficulty passage, with a mean success rate of 87%.

4.3 Gaps and Duplications

No major gaps or duplications were suggested by the analysis of the data on reading. For the most part, university and secondary expectations of student performance were similar, with university teachers being slightly less confident than their secondary counterparts of student achievement in this skill. (The consistent discrepancy parallels the similar consistent difference in the Project III data in teachers' assessments of the level of knowledge of students at the interface. It may be explained by a hypothesis similar to the one advanced in that case--a different perception at the two levels of what constitutes satisfactory performance.) The data suggest that it would be wise to take a future look, however, at four areas--most especially, student capability to deal with reading for inference and implication, and to identify the main idea or purpose of a passage. The other two areas are more properly dealt with under the headings of grammar (what may be an insufficient grasp of reflexive and semi-reflexive verbs) and vocabulary (greater student capability than expected by teachers in the comprehension of idiomatic French.)

5. LISTENING

Although skill in reading has always been an important goal of the secondary school French courses, emphasis on aural comprehension is a relatively recent phenomenon. It was considered by both university and secondary teachers included in Project III to be second in importance only to some aspects of grammar, and equal in importance to reading comprehension (and, in the secondary case, vocabulary). The Project III Listening Test is thus of particular interest to this study.

The Listening Test was composed of five sections, increasingly complex. Each will be discussed in turn.

5.1 Listening Test, Part 1

Part 1 consisted of seven items. The student was supplied with four pictures for each item, and was asked to select the the picture corresponding to the sentence he/she heard on the test tape. Each of these items was given the same sort of assessment by teachers as the items in the first part of the Reading Test (A, B, C or D in the case of secondary teachers; A, B or C in the case of university teachers).

Table 7.5 reports for each item (both in this part of the test and in other parts) the percentage of students correctly answering the item and the percentage of teachers at each level giving various assessments to each item.

All of these items were judged by at least 96% of secondary teachers as falling within the capabilities of their students by the end of Grade Thirteen. Only one item fell below the 96% mark at Grade Thirteen entry. The same item was the only one judged to be old knowledge at university admission by fewer than 94% of university teachers (only 72% classified it A). This item was by far the worst answered, with only 24% of students responding

correctly. Performance was very good (80% or more correct) on four of the items, somewhat lower (52% and 56% respectively) on two other and very low on the one already mentioned. To judge by the response pattern on one poorly answered item, the problem was a failure to distinguish aurally between the words "sous" and "sur"; on the other items the problem was somewhat more complex. The very poorly done item again involved a reflexive verb.

5.2 Listening Test, Part 2

In this section of the test, students heard a brief remark or question. They were then required to select from among four choices printed in their test books the response which would be most likely to be made to this remark or question. The same data are given in Table 7.6 for these items as for those in Part 1.

The close relationship evident in the Reading Test and in the first part of the Listening Test between secondary school assessments and actual student performance on particular items breaks down badly in this section. There is still some correlation between the secondary and university expectations of students leaving Grade Thirteen and entering university, but actual student performance differs strikingly from these expectations, except at the upper extreme. The two items for which the percentage of A and B rankings from secondary teachers was 100% also received the greatest percentages of A rankings from university teachers (94% and 100% respectively), and were the best done (with success rates of 87% and 86%). The same item (number 15) appeared at the bottom of both the secondary and university rankings, with 60% of secondary teachers feeling students should be able to handle this material, and 39% of university teachers agreeing; it was in fact the fifth best done item. The six items which received ratings from secondary teachers which were between 90% and 100% As or Bs had success rates ranging from 52% to 81%, with more of them near the bottom of that range than near the top.

With the exception of item 15, all items were considered within the capabilities of these students by at least 93% of their secondary teachers and 61% of their university teachers. Yet only four items were well answered. (As in the first part of the Reading Test, there is a decided split in the results--four items answered correctly by 79% or more of students, and the other five answered correctly by at most 61% of students.)

There is no obvious similarity among the poorly-answered items which might account for these results (although again one contains a reflexive verb). The assumption for the moment must be that students are considerably less competent in this sort of task than their teachers at either level assume (though here the university expectations come closer to reality). This assumption will be supported or questioned as we proceed to examine the data on similar but more complex tasks.

5.3 Listening Test, Part 3

In Part 3 of the test, the students listened to a series of short dialogues. After each dialogue, they chose from among four printed choices the one statement which was correct according to what they had heard. Data on these items are included in Table 7.6 in the same form as data on the items in the previous parts of the test.

The sharp division into well-done and poorly-done items does not appear in this part of the test; success rates range over the whole interval from 36% to 91%. This is somewhat surprising in view of the consistency of secondary assessments of the items, with five of the eight receiving more than 90% A or B ratings. Those with the smallest percentages of such ratings (70% and 80%) were the worst done; they also received the smallest percentages of A ratings from university teachers (41% and 53%). So the three sets of data correspond at the lower end. However, one item at the very top of both sets of ratings (item 23, with 99% A or B ratings from secondary teachers and 94% A ratings

from university teachers) was relatively poorly done at 56% correct, and in general the spread of success rates was much greater than could have been expected from the inventory data. The results on this section must be regarded as supporting the hypothesis that students are less able in this area than their teachers at both levels believe.

5.4 Listening Test, Part 4

In Part 4 of the test, the students heard a series of short broadcasts or announcements, each followed by a question to which they were to select the appropriate response from among four printed choices. The data on these items are included in Table 7.6.

Student performance on these five items was relatively consistent when compared to that on earlier sections of the test, with all success rates between 51% and 67%. This performance was, however, in some cases less good than might have been expected from the teacher ratings. According to 96% or more of secondary teachers, students should have been competent to deal with items 25 and 26, yet the success rates were fairly low. (These two items were also most widely expected to be old knowledge at the university level.) Performance was poorest on the item which fewest secondary teachers felt that students should be able to deal with.

Again, students seemed in this section to be performing below the expectations of their teachers at the secondary level, though (except in the case of item 29) near the level expected by university teachers.

5.5 Listening Test, Part 5

The last part of the Listening Test consisted of two much longer passages, followed respectively by two and three questions about

the passage. The first was a dialogue between two people; the second involved three voices in a dramatic scene. Data on the five items are in Table 7.6.

Secondary teachers had about the same expectations for both passages, giving no fewer than 89% A or B ratings to all five items. University teachers had somewhat higher expectations of student performance on the first passage. In fact, student performance was considerably better on the second; the mean success rate on its three related items was 77%, as opposed to a mean success rate of 56% on the two items dealing with the first passage.

Performance on the first passage can be described as well below both sets of expectations. That on the second passage was reasonably consistent with secondary teachers' expectations, and as high as or higher than those of university teachers.

5.6 Gaps and Duplications

In spite of the differences that have been pointed out among student performance on various parts of the test, average performance on these parts did not differ greatly. Performance was best on the last part, which included the two lengthy passages; the mean success rate on these items was 70%. The students did most poorly on Part 4--the broadcasts and announcements--with a mean success rate of 60%. Some hypotheses suggest themselves: that students perform better when given a longer passage so that they can pick up clues from the context about the meaning of unfamiliar words and phrases; that the contrast of two or more voices provides additional contextual help to the students; that the double task of recognizing both the spoken words on the tape and the written words in the test book is more difficult in general than that of matching the spoken words with a picture.

But these hypotheses, although they may provide ground for further investigation, are not the main aim of this section. That aim is a comparison of actual test performance with the expectations of both secondary and university teachers about student competence in aural comprehension, and a comparison of those expectations with one another.

If we examine the Listening Test as a whole (see Table 7.7), giving each item equal weight and disregarding the differences between and among parts of the test, it is clear that there is a strong relationship between both sets of expectations and actual performance. In general, as the performance level rose, so did the percentage of teachers expecting students to know the material tested. There are two anomalies--secondary teachers greatly misjudged student knowledge on the four most poorly done items, and university teachers showed less precise judgment of this knowledge on those items in the 51% to 70% correct range. It is the first of these anomalies with which we are concerned; the second is relatively minor.

The four most poorly done items appear to have no common element. Two of the items (numbers 20 and 24) owe their poor success rate largely to the number of students who omitted them (40% and 27% respectively). There may be in these items elements of vocabulary or grammar with which many students are unfamiliar; this suggestion is supported by the relatively low percentages of A and B ratings the items received from their secondary teachers (80% and 70% respectively). However, the percentages of university teachers expecting incoming students to have this knowledge are also relatively low (59% and 41%), so that there cannot be said to be a serious gap here. The other two poorly done items were not omitted by many students; rather, in each case a particular wrong response was heavily selected. This knowledge was assumed by more university teachers (72% and 78%), and there does appear to be a gap with each of these items. It is unfortunately hard to deduce from the items exactly what the nature of this gap is.

If we examine the nine items on which the success rate was over 80%, we find only one for which the percentage of university teachers expecting students to know the material was under 80% (item 32). Thus there does not seem to be any serious problem of duplication.

In general, then, with minor exceptions, student performance on the Listening Test corresponds fairly closely to university expectations of what incoming students should know, and with the assessments of their secondary teachers of the extent of their knowledge. In some parts of the test, performance was somewhat nearer the (generally slightly higher) secondary expectations; in others, it fit more closely with the slightly lower university expectations. But overall, the fit between the knowledge possessed by incoming students and the knowledge expected by their first year university teachers was good.

6. WRITING

The Writing Test consisted of two parts, one of which could be described with some reservations as "objective", and the other of which could be described with some of the same reservations as "essay".

6.1 Writing Test, Part A

This section of the test was concerned primarily with grammar and the ability to apply it in writing. Further discussion will therefore be postponed to the section of this chapter dealing with grammar.

6.2 Writing Test, Part B

The second half of the test dealt much more directly with Topic 5, "Ability to express ideas clearly and correctly in written French". Students were given ten minutes to write a brief composition on a given subject. They were supplied with a number of themes to be included in the composition (the themes were listed in French, thus removing some of the vocabulary factor).

Table 7.8 reports the responses of both secondary and university teachers to four questions about their expectations of student writing performance, the emphasis they give to writing exercises similar to those used in the test, and their assessment of the difficulty of the exercise in the test.

76% of secondary teachers reported that by the end of the Grade Thirteen course, better than three-quarters of their students should be able to produce an acceptable composition of this sort. 66% of university teachers expected that more than three-quarters of their incoming students should be able to perform this task. Almost identical proportions of teachers at the two levels reported giving heavy or moderately heavy emphasis to this type of assignment--57% of secondary teachers and 56% of university teachers. About three-quarters of each group felt that the assignment was at a suitable level of difficulty.

Each writing exercise was marked by two different markers, and assigned three different marks by each. Each marker counted, first, the total number of intelligible clauses in the composition; second, the number of these clauses which were grammatically correct; and third, the number of clauses which were correct in vocabulary.

Each pair of marks was averaged to produce a final set of three marks for the writing exercise--total number of intelligible clauses, number of grammatically correct clauses, and number of clauses correct in vocabulary.

A number of further scores were calculated from these figures; details are given in Appendix A3 to the Project II report. Only two of these scores will be considered here, for reasons outlined in that Appendix. They are (where a = total number of intelligible clauses, b = number of clauses correct in grammar, and c = number of clauses correct in vocabulary):

$$(a) \text{ Writing Quantity} = 3b + 2c$$

$$(b) \text{ Writing Quality} = 40b/a + 30c/a$$

6.3 Gaps and Duplications

Table 7.9 present summary performance data for the writing exercise. The average student produced about 14 clauses in the allotted ten minutes, of which about 11 contained no errors in vocabulary and about seven contained no grammatical errors. This would appear to be acceptable performance in terms of length and use of vocabulary, though somewhat less than desirable in terms of grammar.

What is of most interest and concern, however, is the amount of variability evident in these scores. Of the 54 students tested, six were unable to produce more than seven clauses altogether, but nine wrote more than 20 clauses. Eight students failed to produce as many as six clauses correct in vocabulary, but six others wrote 16 or more such clauses. Ten students managed no more than two grammatically correct clauses; four produced 15 or more clauses without grammatical errors. That this variation in number of correct clauses written is more than a function of variation in total length can be seen by the large range of the Writing Quality scores and the relatively large standard deviation of those scores (which are independent of total length).

This variation cannot help but present problems to teachers at the first year university level. They are receiving on the one hand some students who demonstrate considerable fluency and facility in writing in French; on the other hand, others of their incoming students have little but a rudimentary skill in the written language. This may in part explain the diversity reported in Table 7.8 in the amount of emphasis given to writing skills in university courses. And the variation in performance may well be closely related to the similar diversity of treatment at the Grade Thirteen level, as also reported in that table.

The whole area of writing must, on the basis of these test results, be considered as suffering from both gaps and duplications. Any university writing program aimed at the students less able in writing will be far below the level of the better students, and any program which assumes the high level of skill demonstrated by many of the better students will be beyond the grasp of those who have not mastered the written language.

7. SPEAKING

The Speaking Test was composed of four parts, testing respectively pronunciation; structural control; oral reading; and fluency. Although elements of grammar and vocabulary were involved, the main focus of the test was on Topic 13, "Fluency in spoken French".

7.1 Speaking Test, Part I

Part I of the test required the student to repeat onto a tape 16 brief sentences after hearing them twice from a second tape. Both secondary and university teachers assessed each phrase on the ABCD or ABC scale used with portions of the other tests of French. The assessments are not reported here because of the small amount of variation from item to item. For the 16

sentences, the range in percentages of secondary teachers assessing them as A or B is from 90% to 98%; the range of university teachers assessing them as A is from 94% to 100%. In other words, both groups of teachers were confident that all tasks were within the capabilities of the students.

The total number of points it was possible for a student to earn on this section of the test was 29. The 16 sentences contained 29 critical phonemes or phonological features; the student's rendition of each was scored 1 (acceptable) or 0 (unacceptable). The range of scores assigned was from 13 to 29, with a mean of 21.09 and a standard deviation of 3.39--consistent and satisfactory performance.

7.2 Speaking Test, Part II

This part of the test was designed to assess structural control. It contained ten items, each of which required the student to examine a simple picture, listen to a question about the picture, and make a brief reply in sentence form. Teachers were also asked to assess the content of these items on the ABCD or ABC scale; since there was more variation here than in the first part of the test, their responses are summarized in Table 7.10.

A total of 40 points could be gained for this part of the test. Each response was scored on the following scale:

- 0 - no response; a response of "yes" or "no"; an inappropriate or unintelligible response; a response of "I don't know"
- 1 - an accumulation of serious errors; one serious error accompanied by one or several minor errors
- 2 - one serious error or an accumulation of minor errors

3 - one or two minor errors

4 - completely correct

The actual scores ranged from 0 to 36, with a mean of 18.73 and a standard deviation of 8.48--considerably more variation than was present in the pronunciation part of the test, and an average level of achievement that must be considered less than satisfactory. (The mean of 18.73 represents a mean individual item score of 0.19--at least one serious error or an accumulation of minor errors.) There was only one of these items not considered within the students' competence by at least 90% of secondary teachers, and one other which was expected at university entry by fewer than 75% of university teachers. Judged in the light of these expectations, student performance was disappointing. The wide range of scores (one-fifth of the students had scores above 25, while another fifth had scores below 12) presents the same problems to university teachers as does the similar score range on the writing exercise forming part of the Writing Test. A class containing students of such disparate levels of ability cannot be a simple one to teach satisfactorily.

7.3 Speaking Test, Part III

In the third part of the test, the student was required to read aloud a passage in French after being given three minutes in which to study the text. Table 7.10 reports the ratings by secondary and university teachers of the difficulty of this passage, along with their opinions on what percentage of their students should be able to perform this task acceptably on entry to and on exit from their courses. The teachers also specified the amount of emphasis given in their courses to oral reading.

There was a very large difference of opinion between secondary and university teachers about the ability of students at the interface in oral reading. 89% of secondary teachers stated that on exit from Grade Thirteen courses, more than

three-quarters of students should be able to read the given passage acceptably. However, only 50% of university teachers expected this capability from more than three-quarters of their incoming students. As well, a higher proportion of secondary teachers judged the passage to be very easy or somewhat easy--27% in contrast to 11% of university teachers. These contrasts may indicate that teachers at the two level attached different meanings to the word "acceptable".

A total of 14 points was allotted to this part of the test. The marker considered 24 points concerning particulars of pronunciation (including stress and intonation), and assigned each a mark of 1 or 0. Scores varied from 7 to 23, with a mean of 15.71 and a standard deviation of 3.85. As for the pronunciation portion of the test, results were consistent, and satisfactory though not excellent.

7.4 Speaking Test, Part IV

The last part of the Speaking Test was designed to test fluency in spoken French. Students were presented with two similar tasks. In the first, they were asked to choose among three sets of three pictures, each in "comic strip" format--i.e., telling a connected story--and, after one minute for preparation, to describe in 30 seconds what was happening in the pictures. In the second task, the choice was among three single pictures; after one minute of preparation, the student was to describe in 30 seconds the probable events that took place before the events in the picture, the actual events, and the probable outcome of the situation.

Table 7.12 gives the comments of teachers on the difficulty of each of these tasks, and the percentage of their students that they felt should be able to perform the tasks acceptably on entry to and exit from their courses. It also includes a description of the amount of emphasis teachers reported giving to the development of student ability in such tasks.

For the first task, 74% of secondary teachers felt the exercise was within the capabilities of more than three-quarters of their departing students. Only 59% of university teachers, however, expected more than three-quarters of their students to arrive with this capability. The secondary figure dropped slightly to 67% for the second task; the university figure was again 59%, although fewer of these than in the former case expected all students to be able to perform this task acceptably. Both tasks were judged to be at a reasonable level of difficulty by both secondary and university teachers. The development of student competence in this sort of task was reported as much more heavily emphasized at university than in Grade Thirteen. Over 80% of university teachers stated that they gave it heavy or moderately heavy emphasis, while only 34% of secondary teachers stressed it to this degree and 18% gave it no emphasis at all.

Student performance on these tasks was scored on a number of scales. One each of the two tasks, a student was assigned six scores:

- (a) total number of intelligible clauses used;
- (b) number of clauses differing in grammatical structure;
- (c) number of clauses correct in structure;
- (d) number of clauses correct in morphology;
- (e) number of clauses correct in vocabulary; and
- (f) number of clauses correct in pronunciation.

These scores were then added over the two tasks to produce a third set of six scores. A further score of from 0 to 4 was assigned on a global basis to performance on the entire fluency task, with the points defined as:

0 - very bad; or no response

1 - poor but passable

2 - satisfactory

3 - good

4 - very good

As for the writing exercise discussed earlier, some further scores were computed from the ones listed above, using the scores for the combined task. The two of interest here are:

$$\text{Fluency Quantity} = 2b + 2c + d + 2e = 2f$$

$$\text{Fluency Quality} = (6c + 3d + 6e + 6f)/a$$

with both scores set to 0 in the case where $a = 0$.

The theoretical range of the Fluency Quality score is from 0 to 21. The Fluency Quantity score has a lower limit of 0, and no theoretical upper limit; in the case of this administration, the upper limit would be 216, since the greatest number of clauses produced by any student was 24.

Table 7.12 reports summary data on student performance on the fluency tasks. The first task appears to have been slightly easier for the students, who produced somewhat longer responses to this task with a correspondingly higher production of clauses correct in structure, morphology, vocabulary and pronunciation (although the number of different grammatical structures used was slightly fewer than for the second task).

In the 60 seconds allotted for the two exercises, the average student produced about 13 clauses, of which about seven were correct in morphology, nine correct in vocabulary, nine correct in pronunciation, and 10 correct in structure. About nine

different grammatical structures were employed. The mean global fluency rating was 1.52--halfway between "poor but passable" and "satisfactory". Where these results can be compared to results on the writing exercise, they are similar--for example, the ratio of mean number of clauses correct in vocabulary to mean total number of clauses was 0.73 for the fluency task, 0.77 for the writing exercise. More clauses were faulty in morphology than in any other respect--over 40% of all clauses produced had some morphological error, while, for example, fewer than one-quarter of all clauses contained an error in structure.

What is apparent here, as in the writing exercise, is the very large variability in performance. Combined responses ranged in length from five to 24 clauses; the number of different grammatical structures employed varied from three to 18. Similar variation appeared in the numbers of clauses correct in the various respects of interest to the study. Of 54 students assigned a global fluency rating, eight were considered "very bad"; eight others were rated "good" or "very good". The large range of Fluency Quality scores indicates that correctness of expression was not closely tied to the length of the answer.

7.5 Gaps and Duplications

Pronunciation was tested in Parts I, III and IV of this test. The test results in Part I, in which students were to repeat sentences after hearing them spoken, indicated satisfactory performance, with relatively little variation. Students did only slightly less well in Part III, which tested oral reading; again results were fairly consistent. Average results were at about the same level in Part IV (Fluency), but it is not possible to measure the extent of variation here because separate scores were not computed for the ratio of clauses correct in pronunciation to total clauses.

There does not appear to be either a sizeable gap or a sizeable duplication in the area of pronunciation. University teachers reported that they felt students should perform competently on Part I, and the students did so. They felt that the oral reading passage was slightly on the difficult side, and student performance here was slightly less good. Teacher expectations about the level of pronunciation skills in the fluency task were not elicited, nor were deductions made from the performance data about performance in pronunciation as a separate issue; therefore no conclusions about gaps and duplications in pronunciation can be drawn from Part IV.

Structural control was measured in Parts II and IV of the test; again, the data do not permit an examination of this as a separate issue in Part IV. The Part II test results were unsatisfactory; on the average, students appear to have committed at least one serious error or several minor errors in each brief response. For each item, between 72% and 100% of university teachers felt that students entering their courses should have the knowledge required to respond satisfactorily to the item.

Overall, this must be considered a gap area. This assessment must, however, be qualified by a consideration of the very high degree of variability in student performance pointed out earlier. University courses based on the assumption that incoming students are performing at the level indicated by the mean score on this section would be grossly ill-suited to the sizeable number of students performing very well in this area.

This same degree of variability appears in the fluency score, where there were large differences in both the length and the quality of the responses produced by various students. Some students had Fluency Quality scores approaching perfection (19.93 of a possible 21) and/or global fluency ratings of 4 out of 4. Others had very low Fluency Quality scores and/or global ratings of 0. And the proportion of students at or near the extremes of these ranges was large. The same diagnosis of simultaneous gap and duplication must be made here.

8. GRAMMAR

A command of grammar is only incidentally measured in the Reading Test and Listening Test, in that success or failure on a particular item may be determined by the student's awareness of a grammatical distinction which affects the meaning of a phrase or sentence. There is little that can be deduced from the data on those tests about the students' grammatical competence. This competence is measured much more directly in parts of both the Speaking Test and the Writing Test.

The two sections of the Speaking Test concerned with grammar to any extent have been discussed at length above. To recapitulate: performance on the section of that test dealing with structural control was poor, with on the average each spoken sentence containing at least one serious error or several minor errors. However, variability was very great. In the fluency portion of the test, the most frequently committed errors were those of morphology and the rarest those of structure. Variability could not be estimated accurately for grammar alone, but the Writing Quality score was assigned in a way which made it more heavily dependent on correctness in structure and morphology than on correctness in either vocabulary or pronunciation (though less dependent on the grammatical factors than on a combination of vocabulary and pronunciation).

What remains to be examined here is Part A of the Writing Test, which is concerned almost entirely with grammar. It consisted of two groups of incomplete sentences. The first 26 sentences contained one blank each, which the student was to fill with what he/she considered to be the correct word. In each case, one and only one word was considered acceptable, and the student received a mark for the word only if it was spelled correctly, including accents. The remaining six sentences contained two or three blanks each, and were accompanied by a "pattern" sentence with certain words underlined. These underlined words were to be inserted in the blanks in the main sentence, with all appropriate changes in number, gender and/or person to fit the new context.

Again, a mark was allotted for each blank only if the inserted word was correct in all respects. A total of 42 marks was available, one for each blank.

A count was done of the number of these items judged to test each of Topics 1, 2 and 3 (basic morphology and syntax; conjugation of all regular and auxiliary verbs; and common tenses and moods). A number of items tested more than one of these topics, so that the total for the three topics substantially exceeds 32. 26 of the items dealt with basic morphology and syntax, nine with verb conjugation, and five with common tenses and moods (these five are all included in the nine on verb conjugation).

A more precise indication of the content of this test section may be given by a tally of the number of each different part of speech required to fill the 42 blanks. The words to be supplied for the first 26 sentences included nine pronouns, two adjectives (one demonstrative, one interrogative), three prepositions, seven verbs or parts of verbs, two adverbs, and three negative particles; each of these words was uniquely determined by the context. The 16 blanks in the remaining sentences were to be filled by two nouns, one indefinite pronoun, seven adjectives, one article, and five verbs or parts of verbs; the emphasis in these sentences was on agreement in gender, number and person and corresponding morphological changes.

Secondary and university teachers assessed these items on the ABCD and ABC scales discussed earlier; Table 7.14 reports their responses. All items received at least 87% A or B ratings from secondary teachers; all but seven received at least 75% A ratings from university teachers, with three of these seven at the 72% mark. It would therefore seem reasonable to expect good student performance on this test section.

Table 7.15 reports the distribution of student scores on this part of the test. They are much lower than could have been predicted from the teacher assessments, with students on the

average scoring only about one-half the possible points. Variability is again large.

This section, then, shows the same pattern as the portion of the Speaking Test in structural control--poor and highly variable performance. This a further "gap and duplication" situation, but clearly the gap problem is the more serious; relatively fewer students evidenced excellence in grammar, than was the case with other skills, and the variability, although great, was centred about a lower level of competence.

9. SUMMARY

There was a generally good fit between teacher expectations and student performance in both reading and listening skills; minor discrepancies are noted in the sections of this chapter dealing with those skills. Student performance was consistently at an acceptable level.

This pattern is also true of pronunciation, but a different pattern emerges when it comes to other skills connected with speaking and with writing. Although average performance in writing and speaking was at a reasonable level, the amount of variability in performance was very great, ranging from almost total inability to speak or to write fluently to a very high degree of facility in using the language. In these areas one can only say that there exist simultaneously a gap and a duplication. The large number of students performing poorly cannot help but be out of their depth in a university course; the equally large number performing well must be frustrated by the need for university teachers to devote time to bringing others up to the same standard of performance.

Somewhat the same situation holds with respect to grammar; variability was great here as well. However, the average standard of performance was lower here than would be suggested by the

writing and speaking results as a whole, and consequently this is more properly defined as a gap, with some accompanying problem of duplication for those students who have achieved a high level of mastery of morphology and syntax.

In summary, then, the "passive" skills of reading and listening were handled satisfactorily and with relatively little variability; performance largely matched teacher expectations. The "active" skills of writing and speaking were problem areas (except for pronunciation). Average performance on writing and speaking, excluding the component of grammatical correctness, was adequate, but with very high variability. Average command of grammar was not adequate, and the same degree of variability was evidenced.

TABLE 7.1
PROJECT III LIST OF TOPICS

1. Grammar--basic morphology and syntax
2. Grammar--conjugation of all regular and auxiliary verbs
3. Grammar--common tenses and moods
4. Ability to apply grammar in writing and speaking
5. Ability to express ideas clearly and correctly in written French
6. Skill in translation--English to French
7. Skill in translation--French to English
8. Vocabulary--general
9. Vocabulary--idiomatic expressions
10. Vocabulary--Canadianisms
11. Reading comprehension
12. Aural comprehension
13. Fluency in spoken French
14. Literary history
15. Concepts of literary criticism
16. Vocabulary of literary criticism
17. Understanding of cultural differences and similarities

TABLE 7.2

ASSESSMENT BY TEACHERS OF STUDENT KNOWLEDGE
AT THE INTERFACE

<u>Topic</u>	<u>Mean reported secondary exit level* (and rank)</u>		<u>Mean reported university entrance level* (and rank)</u>	
1	2.7	(6)	1.6	(4)
2	3.0	(1)	1.7	(1)
3	2.8	(2)	1.6	(4)
4	2.6	(7)	1.6	(4)
5	2.5	(9)	1.1	(10)
6	2.3	(12)	1.1	(10)
7	2.3	(12)	1.4	(7)
8	2.8	(2)	1.4	(7)
9	2.5	(9)	1.0	(12)
10	1.2	(14)	0.7	(14)
11	2.8	(2)	1.7	(1)
12	2.8	(2)	1.7	(1)
13	2.6	(7)	1.4	(7)
14	1.2	(14)	0.3	(17)
15	1.2	(14)	0.5	(15)
16	1.0	(17)	0.5	(15)
17	2.4	(11)	1.0	(12)

*RESPONSE KEY:

- 0 - No competence
- 1 - Minimal competence
- 2 - Moderate competence
- 3 - Competence in varied situations, some originality
- 4 - Mastery, competence in high level, creative situations

TABLE 7.3
SUMMARY DATA: PART A OF READING TEST

<u>Item</u>	<u>% correct</u>	<u>Secondary assessment*</u>				<u>University assessment*</u>		
		<u>A</u>	<u>B</u>	<u>C</u>	<u>D</u>	<u>A</u>	<u>B</u>	<u>C</u>
1	97	100	0	0	0	94	6	0
2	81	78	12	8	2	78	17	6
3	62	98	2	0	0	94	6	0
4	78	98	2	0	0	100	0	0
5	69	92	6	2	0	78	22	0
6	41	37	41	16	6	50	44	6
7	83	50	23	20	6	61	28	11
8	49	16	43	27	14	33	61	6
9	31	38	25	27	10	39	50	11
10	32	33	41	22	5	67	28	6
11	63	23	31	23	22	50	50	0
12	78	62	23	12	2	89	11	0
13	60	56	28	14	2	67	33	0
14	38	29	19	32	19	50	39	11
15	45	27	31	22	20	50	44	6
16	13	20	38	22	20	56	22	22
17	26	11	27	32	30	28	67	6

*See text, Subsection 4.1, for explanation.

TABLE 7.4

TEACHER RESPONSES TO QUESTIONS ABOUT READING TEST, PART B,
AND READING COMPREHENSION SKILLS

	<u>Secondary teachers(%)</u>	<u>University teachers(%)</u>
How many of the students entering courses at (your) level should have the ability to read passages like those in the test for <u>literal understanding</u> ?		
100%	24	50
76%-99%	33	28
51%-75%	22	0
26%-50%	8	22
1%-25%	0	0
None	0	0
What emphasis is given (in your course) to <u>reading for literal understanding</u> ?		
Heavy emphasis	36	44
Moderately heavy emphasis	34	22
Light emphasis	25	17
Individual remediation only	5	11
None	0	6
How many students who successfully <u>complete</u> (your) course should be able to read passages like those in the test for <u>literal understanding</u> ?		
100%	50	67
76%-99%	38	28
51%-75%	8	0
26%-50%	3	6
1%-25%	0	0
None	2	0

TABLE 7.4 (continued)

	<u>Secondary teachers(%)</u>	<u>University teachers(%)</u>
How many of the students <u>entering</u> courses at (your) level should be able to <u>identify the main idea</u> <u>or purpose</u> of passages like those in the test?		
100%	25	44
76%-99%	35	17
51%-75%	27	22
26%-50%	10	11
1%-25%	3	6
None	0	0
What emphasis is given (in your course) to <u>reading to identify</u> <u>the main idea or purpose</u> ?		
Heavy emphasis	48	39
Moderately heavy emphasis	42	56
Light emphasis	8	6
Individual remediation only	2	0
None	0	0
How many students who successfully <u>complete</u> (your) course should be able to <u>identify the main idea or</u> <u>purpose</u> of passages like those in the test?		
100%	61	67
76%-99%	36	28
51%-75%	2	0
26%-50%	2	6
1%-25%	0	0
None	0	0

TABLE 7.4 (continued)

	<u>Secondary teachers(%)</u>	<u>University teachers(%)</u>
How many of the students entering courses at (your) level should be able to <u>draw inferences from, and</u> <u>see implications in, the ideas</u> expressed in passages like those in the test?		
100%	8	6
76%-99%	24	39
51%-75%	30	17
26%-50%	19	22
1%-25%	19	11
None	0	6
What emphasis is given (in your course) to <u>drawing inferences</u> <u>and seeing implications?</u>		
Heavy emphasis	27	39
Moderately heavy emphasis	50	33
Light emphasis	19	22
Individual remediation only	3	6
None	2	0
How many students who successfully <u>complete</u> (your) course should be able to <u>draw inferences and see</u> <u>implications?</u>		
100%	16	22
76%-99%	38	33
51%-75%	30	22
26%-50%	14	17
1%-25%	2	6
None	0	0

TABLE 7.5

TEACHER ASSESSMENTS OF DIFFICULTY, AND STUDENT
PERFORMANCE DATA, PART B OF READING TEST

<u>Passage</u>	<u>Mean secondary assessment of difficulty*</u>	<u>Mean university assessment of difficulty*</u>	<u>%correct for related items</u>
1	1.8	2.1	Item 18: 95 Item 19: 88 Item 20: 82 Item 21: 97 Mean: 91
2	1.8	2.2	Item 22: 85 Item 23: 85 Item 24: 90 Mean: 87
3	2.8	3.3	Item 25: 81 Item 26: 54 Item 27: 73 Mean: 69
4	2.6	2.9	Item 28: 89 Item 29: 87 Item 30: 86 Mean: 87
5	3.4	3.8	Item 31: 48 Item 32: 62 Item 33: 24 Item 34: 30 Mean: 41
6	3.4	3.7	Item 35: 66 Item 36: 85 Item 37: 78 Item 38: 32 Item 39: 52 Mean: 63

*See text, subsection 4.2, for explanation of scale.

TABLE 7.6
SUMMARY DATA: LISTENING TEST

Item	% correct	A	Secondary assessment*		D	University assessment*		
			B	C		A	B	C
Part 1								
1	56	100	0	0	0	100	0	0
2	80	96	4	0	0	100	0	0
3	85	96	0	4	0	100	0	0
4	86	98	0	0	2	100	0	0
5	89	100	0	0	0	100	0	0
6	24	82	16	2	0	72	17	11
7	52	98	2	0	0	94	0	6
Part 2								
8	87	100	0	0	0	94	0	6
9	86	100	0	0	0	100	0	0
10	81	89	5	4	2	83	17	0
11	52	93	5	2	0	83	11	6
12	53	71	25	4	0	78	22	0
13	54	73	20	7	0	67	33	0
14	79	75	18	5	2	61	28	11
15	61	27	33	18	22	39	50	11
16	52	80	15	5	0	61	33	6
Part 3								
17	91	96	2	2	0	94	0	6
18	71	94	4	2	0	88	6	6
19	82	94	4	2	0	94	0	6
20	36	56	24	17	4	53	35	12
21	68	65	20	7	7	59	41	0
22	68	78	17	2	4	82	12	6
23	56	93	6	2	0	94	6	0
24	43	48	22	17	13	41	35	24
Part 4								
25	61	65	33	2	0	78	17	6
26	67	89	7	4	0	83	11	6
27	59	33	43	17	7	56	39	6
28	63	53	34	9	4	56	44	0
29	51	29	40	19	12	67	28	6

TABLE 7.6 (continued)

Item	% correct	A	Secondary assessment*		D	University assessment*		
			B	C		A	B	C
Part 5								
30	63	87	6	6	2	83	11	6
31	49	87	8	6	0	78	6	6
32	87	81	8	11	0	72	11	17
33	70	83	10	8	0	61	17	22
34	79	83	12	6	0	72	11	17

*See text, Subsection 4.1, for explanation.

TABLE 7.7
COMPARISON OF STUDENT PERFORMANCE ON LISTENING TEST
WITH TEACHER ASSESSMENTS

<u>% correct</u>	<u>No. of items</u>	<u>Mean % of A or B ratings by secondary teachers*</u>	<u>Mean % of A ratings by university teachers*</u>
0- 50	4	86	61
51- 60	9	74	78
61- 70	8	88	68
71- 80	4	96	80
81-100	9	97	93

*A - old knowledge

B - new knowledge, taught to all students

TABLE 7.8

TEACHER RESPONSES TO QUESTIONS ABOUT WRITING TEST
AND WRITING SKILLS

	<u>Secondary teachers(%)</u>	<u>University teachers(%)</u>
How many of the students <u>entering</u> courses (at your level) should be able to write an acceptable composition of this type?		
100%	28	33
76%-99%	32	33
51%-75%	17	22
26%-50%	13	6
1%-25%	10	6
None	0	0
How much emphasis do you give in your course to developing student competence in writing French compositions of this type?		
Heavy emphasis	17	17
Moderately heavy emphasis	40	39
Light emphasis	40	28
Individual remediation only	2	11
No emphasis	2	6
How many of the students who successfully <u>complete</u> (your) course should be able to write an acceptable composition of this kind?		
100%	48	56
76%-99%	28	39
51%-75%	18	0
26%-50%	5	6
1%-25%	2	0
None	0	0
Was the writing assignment in Part B at a reasonable level of difficulty for students (at this level) in French?		
Yes	72	78
No	28	22

TABLE 7.9
SUMMARY DATA ON STUDENT PERFORMANCE ON
WRITING TEST, PART B

<u>Score</u>	<u>Mean</u>	<u>S.D.</u>	<u>Range</u>
Total intelligible clauses	14.09	5.56	1.50 - 27.00
Total clauses correct in grammar	6.74	4.42	0.50 - 20.50
Total clauses correct in vocabulary	10.79	4.74	0.50 - 26.00
Writing Quantity	41.80	21.61	2.50 - 110.50
Writing Quality	40.50	9.10	23.04 - 59.39

TABLE 7.10
TEACHER ASSESSMENTS OF ITEMS
IN PART II OF SPEAKING TEST

<u>Item*</u>	Secondary assessment(%)**				University assessment(%)**		
	<u>A</u>	<u>B</u>	<u>C</u>	<u>D</u>	<u>A</u>	<u>B</u>	<u>C</u>
3	98	2	0	0	78	22	0
4	100	0	0	0	100	0	0
5	79	12	9	2	72	11	17
6	70	14	11	5	78	17	6
7	100	0	0	0	89	11	0
8	70	21	5	4	89	11	0
9	84	12	4	0	78	17	6
10	96	4	0	0	100	0	0
11	100	0	0	0	100	0	0
12	77	16	5	2	89	11	0

*Items 1 and 2 were practice items.

**See text, Subsection 4.1, for explanation of assessments.

TABLE 7.11

TEACHER RESPONSES TO QUESTIONS ABOUT SPEAKING TEST,
PART III, AND ORAL READING SKILLS

	<u>Secondary teachers(%)</u>	<u>University teachers(%)</u>
What is your estimate of the reading difficulty of this passage for students (at this level)?		
Very easy	3	0
Somewhat easy	24	11
About right	63	67
Somewhat difficult	10	11
Very difficult	0	11
How many of the students entering a course (at your level) should be able to give an acceptable reading of this passage?		
100%	29	17
76%-99%	36	33
51%-75%	27	39
26%-50%	8	6
1%-25%	0	6
None	0	0
How much emphasis do you give in your teaching to developing student competence in reading prose passages of this kind out loud?		
Heavy emphasis	20	0
Moderately heavy emphasis	36	39
Light emphasis	34	50
Individual remediation only	10	6
No emphasis	0	6

TABLE 7.11 (continued)

	<u>Secondary teachers(%)</u>	<u>University teachers(%)</u>
How many of the students who successfully complete (your) course should be able to give an acceptable reading of this passage?		
100%	47	24
76%-99%	42	76
51%-75%	10	0
26%-50%	0	0
1%-25%	0	0
None	0	0

TABLE 7.12

TEACHER RESPONSES TO QUESTIONS ABOUT SPEAKING TEST,
PART IV, AND FLUENCY SKILLS

	<u>Secondary teachers(%)</u>	<u>University teachers(%)</u>
How many of the students <u>entering</u> a course (at your level) should be able to give an acceptable response to the <u>first</u> fluency task?		
100%	18	24
76%-99%	32	35
51%-75%	27	24
26%-50%	16	6
1%-25%	7	12
None	0	0
. . . the <u>second</u> fluency task?		
100%	12	12
76%-99%	28	47
51%-75%	30	29
26%-50%	16	6
1%-25%	14	6
None	0	0
How much emphasis do you give in your teaching to the development of student ability to adequately handle this kind of speaking task?		
Heavy emphasis	11	29
Moderately heavy emphasis	23	53
Light emphasis	45	18
Individual remediation only	4	0
No emphasis	18	0

TABLE 7.12 (continued)

	<u>Secondary teachers(%)</u>	<u>University teachers(%)</u>
How many of the students who successfully <u>complete</u> your course should be able to give an acceptable response to the <u>first</u> fluency task?		
100%	33	62
76%-99%	41	25
51%-75%	19	6
26%-50%	7	6
1%-25%	0	0
None	0	0
. . . the <u>second</u> fluency task?		
100%	39	44
76%-99%	28	38
51%-75%	23	12
26%-50%	5	6
1%-25%	5	0
None	0	0
Was the first fluency task at a reasonable level of difficulty for students (at this level)?		
Yes	84	75
No	16	25
Was the second fluency task at a reasonable level of difficulty for students (at this level)?		
Yes	86	94
No	14	6

TABLE 7.13
SUMMARY DATA ON STUDENT PERFORMANCE ON
SPEAKING TEST, PART IV

<u>Score</u>		<u>Task I</u>	<u>Task II</u>	<u>Combined Task*</u>
Total clauses:	Mean	6.54	6.08	12.76
	S.D.	2.22	2.40	4.13
	Low score	2.00	2.00	5.00
	High Score	14.00	10.00	24.00
No. of different structures:	Mean	4.27	4.33	8.64
	S.D.	1.93	1.84	3.43
	Low score	2.00	1.00	3.00
	High score	10.00	9.00	18.00
Clauses correct in structure:	Mean	5.02	4.73	9.80
	S.D.	2.34	2.34	4.21
	Low score	0.00	1.00	1.00
	High score	12.00	9.00	21.00
Clauses correct in morphology:	Mean	4.08	3.14	7.28
	S.D.	2.94	2.04	4.58
	Low score	0.00	0.00	0.00
	High score	12.00	9.00	19.00
Clauses correct in vocabulary:	Mean	4.71	4.43	9.28
	S.D.	2.55	2.38	4.41
	Low score	1.00	0.00	1.00
	High score	12.00	10.00	21.00
Clauses correct in pronunciation:	Mean	4.75	4.63	9.48
	S.D.	2.60	2.57	4.71
	Low score	0.00	1.00	1.00
	High score	12.00	10.00	22.00
Fluency Quantity:	Mean			81.68
	S.D.			35.34
	Low score			18.00
	High score			182.00
Fluency Quality:	Mean			14.66
	S.D.			3.28
	Low score			5.14
	High score			19.93

TABLE 7.13 (continued)

<u>Score</u>	<u>Task I</u>	<u>Task II</u>	<u>Combined Task*</u>
Global fluency rating:			
Mean			1.52
S.D.			1.09
Low score			0.00
High score			4.00

*Figures for the combined task do not correspond exactly to those for the separate tasks because a few students completed only one of the two. Combined task calculations were carried out only for students who did both.

TABLE 7.14
TEACHER ASSESSMENTS OF ITEMS
IN PART A OF WRITING TEST

Item	Secondary assessment(%)*				University assessment(%)*		
	A	B	C	D	A	B	C
1	93	3	3	0	94	6	0
2	97	2	2	0	89	11	0
3	93	5	2	0	78	22	0
4	85	14	2	0	100	0	0
5	58	33	7	2	83	11	6
6	82	13	3	2	89	11	0
7	85	12	3	0	83	17	0
8	60	27	8	5	56	33	11
9	73	20	7	0	67	33	0
10	62	33	2	3	50	44	6
11	90	8	2	0	100	0	0
12	53	46	2	0	72	22	6
13	98	0	2	0	100	0	0
14	62	31	3	3	56	33	11
15	87	12	0	2	89	11	0
16	80	17	3	0	72	28	0
17	93	5	2	0	94	6	0
18	95	3	2	0	83	17	0
19	97	0	3	0	89	11	0
20	97	3	0	0	94	6	0
21	97	2	2	0	100	0	0
22	95	5	0	0	83	17	0
23	97	2	2	0	89	11	0
24	95	5	0	0	94	6	0
25	97	3	0	0	94	6	0
26	77	18	2	3	72	28	0
27	97	2	2	0	89	11	0
28	95	3	2	0	94	6	0
29	90	10	0	0	94	6	0
30	97	3	0	0	100	0	0
31	85	15	0	0	83	17	0
32	97	3	0	0	94	6	0

*See text, Subsection 4.1, for explanation of assessments.

TABLE 7.15
DISTRIBUTION OF SCORES ON
WRITING TEST, PART 1

<u>Score range</u>	<u>No. of students</u>
0.5 - 7.0	2
7.5 - 14.0	10
14.5 - 21.0	20
21.5 - 28.0	12
28.5 - 35.0	9
35.5 - 42.0	1

Mean score = 20.80

Standard deviation = 7.51

CHAPITRE HUIT

FRANCAIS*

1. INTRODUCTION

A beaucoup d'égards la situation des écoles francophones en Ontario en ce qui concerne l'enseignement du français est analogue à celle des écoles anglophones en ce qui concerne l'enseignement de l'anglais. Dans les dernières années du cycle secondaire, la diversité abonde. Les étudiants de 13^{ème} année ne sont pas tenus s'inscrire dans aucun cours de français. Au niveau de la 12^{ème} année, quelques-uns sont inscrits dans les cours de base, quelques-uns dans les cours généraux, quelques-uns dans les cours avancés. Et la matière enseignée dans ces cours varie beaucoup. Les étudiants peuvent apprendre la littérature, ou les média, ou d'autres sujets extrêmement variés regroupés sous le titre "français".

Les problèmes créés par cette diversité furent considérables pour l'équipe du projet II ainsi que pour celle du projet III. La plupart n'ont pas besoin d'être discutés ici de façon exhaustive, mais l'un d'eux fut d'une importance capitale : celui de la sélection des instruments nécessaires aux tests, adaptés à la mesure des compétences linguistiques de ce groupe hétérogène d'étudiants. On devait trouver ou créer des tests assez recherchés

*(Ce chapitre a été préparé avec l'aide de Lilian Nygren, assistante diplômée au Centre des langues modernes, du Département de "Curriculum" O.I.S.E.)

pour opérer une distinction parmi les étudiants en français de 13ème année mais pas trop difficile cependant de façon à pouvoir convenir en totalité pour le test des étudiants de 12ème année du cours général ou du cours de base.

Le comité de sélection des tests du projet II avait bien compris dès le début qu'un tel test ne pouvait pas être lié au programme. Il existait non seulement parmi les étudiants qui devaient passer les tests, la grande différence de niveau mentionnée ci-dessus, mais encore une trop grande variété dans le contenu des cours à un niveau particulier. Il y avait également un manque de tests appropriés dans beaucoup de disciplines : par exemple, littérature et expression orale. Il ne serait possible que de tester certains secteurs de base de la connaissance de la langue--compréhension de textes, grammaire et travaux écrits. (Les contraintes de temps inhérentes au projet n'ont pas permis d'élaborer et de faire les validations qui s'imposaient pour de nouveaux tests relatifs aux connaissances de l'étudiant, excepté dans le domaine relativement précis des travaux écrits.

2. LES DONNEES

L'équipe du projet III a mis au point un instrument intitulé Questionnaire sur les descriptions de cours, élaboré pour permettre une description détaillé du contenu des cours des professeurs à la fois du secondaire et du post-secondaire, le temps passé sur les différents aspects du français, l'accent mis sur les différents objectifs ainsi que d'autres données similaires. Les réponses sont rapportées en détail dans le rapport du projet III. Un élément important relatif à ce questionnaire fut l'introduction d'une liste d'objectifs possibles d'un cours en français pour la(les) dernière(s) année(s) du cycle secondaire et la première année du cycle post-secondaire. Nous avons adjoint cette liste ici comme tableau 8.1. Les professeurs devaient évaluer sur une échelle allant de 0 à 7, les compétences de leurs étudiants entrants et sortants en utilisant chacun de ces

objectifs. On leur avait également demandé de préciser le niveau souhaité pour leurs étudiants entrants et le niveau--compte tenu de celui préféré à l'entrée--auquel ils estimaient pouvoir les mener à la fin du cours. Dans ce rapport, les seuls chiffres considérés seront par les niveaux réels d'entrée indiqués par les professeurs du secondaire; bien que les niveaux réels d'entrée indiqués par les professeurs du post-secondaire auraient été intéressants, les réponses à ce niveau ayant été si minces qu'elles sont peu fiables.

Les données du projet II sont de deux sortes. Il y a tout d'abord le rapport des résultats des étudiants aux deux tests passés (ces tests seront décrits plus en détail ci-dessous). En second, les professeurs des étudiants testés ont rempli les tests-inventaires estimatifs qui accompagnaient les tests, et on leur a demandé de faire leurs commentaires à la fois sur les données particulières des tests, leurs propres façons d'enseigner et leurs prévisions dans les secteurs couverts par les tests. (Les tests-inventaires estimatifs ont également été distribués à un petit nombre d'instructeurs du post-secondaire, mais nous n'avons pas reçu de formules remplies; ainsi, aucune donnée de cette sorte n'est disponible au niveau post-secondaire.)

Dans le cas du français--en fait, dans le cas de presque tous les tests de langue utilisés dans le projet II--la concordance des données des deux projets est négligeable. Ceci est mis en évidence dans le tableau 8.1 représentant la liste des objectifs préparée par le projet III. Un astérisque (*) après les objectifs de la liste indique s'ils ont été testés par le projet II--quel que soit le test. La proportion des objectifs couverts par les tests est très faible. Les secteurs entiers de la linguistique et de la culture sont omis; l'expression orale n'est pas testée; la littérature n'est testée que par des réponses à des questions sur un texte écrit, ce qui est loin d'être considéré comme une explication de texte écrite.

De la même façon, l'élaboration des tests ne permet pas de séparer l'examen de quelques-unes des compétences testées. L'orthographe, par exemple, n'est testée que dans le contexte de l'orthographe d'accord, que l'on peut plus vraisemblablement considérer comme grammaire. La stylistique est des plus importantes en référence au test de composition écrite, mais n'a dans aucune mesure été examinée en tant que question séparée dans l'attribution de notes à ce test. Etant donné ces limites, il y a peu d'intérêt à tenter de faire concorder les résultats des tests avec la liste des objectifs en aucune façon à moins de s'en tenir à une vue générale. Les tests devront être considérés comme mesurant des compétences linguistiques fondamentales plutôt que comme tests du contenu des cours de français pour les 12ème et 13ème années (test étant compris dans un sens large).

A l'intérieur de ce domaine dont nous acceptons les limites, les tests traitent de ce qui semble être des objectifs importants des cours de français. Cette constatation est appuyée par l'évidence qui se dégage des tests-inventaires estimatifs dont on fera le rapport tout au long de ce chapitre.

3. PROBLEMES D'INTERPRETATION DES DONNEES

Quelques-uns des problèmes d'interprétation ont été suggérés ci-dessus--en particulier la concordance négligeable entre les données des deux projets et la couverture limitée du contenu du programme fourni par les tests. Trois autres problèmes seront discutés ici.

Le premier est commun à tous les tests standardisés et est discuté de façon plus détaillée au chapitre 2 de ce rapport. Le test de compréhension en lecture et de connaissance de la langue, comme les autres tests standardisés, a été élaboré de façon à maximiser la distinction entre les étudiants. En général, pour quelqu'un qui n'est pas familiarisé avec les tests standardisés, la répartition des points pour un tel test semble indiquer que

ceux qui l'ont passé ont une maigre connaissance de son contenu. Les étudiants les moins doués ont ce qui apparaît être des résultats désastreusement bas, et même l'étudiant moyen obtient un résultat qui pourrait simplement être considéré comme "passable" dans un test normal en classe. Il est nécessaire de se garder d'une interprétation comme celle-ci. C'est spécialement important dans ce cas, quand on examine les résultats des étudiants de 12ème année, puisque--comme on le verra plus tard--les sections de ce test faisaient à l'origine partie d'un test destiné à des étudiants en dernière année du secondaire qui se préparaient à entrer à l'université, et difficile même pour eux. Il aurait été très surprenant que les étudiants de 12ème année obtiennent de bons résultats à ce test.

Un autre facteur qui crée des difficultés dans l'interprétation des résultats est que (en parallèle au test de l'anglais dans les écoles anglophones) on n'a pas fait passer le test comme un ensemble. Il a été partagé en deux formules à peu près équivalentes (formule 1 et formule 2) comprenant chacune environ la moitié des éléments de chaque type inclus dans le test. Certains étudiants ont passé les deux formules ; d'autres n'ont passé que la formule 1 ; d'autres encore que la formule 2. En pratique, la formule 2 s'est avérée être considérablement plus facile que la formule 1. Ceci n'est pas un problème dont nous ayons à tenir compte dans cette analyse dont l'objet est le succès aux éléments de tests individuels, mais il en devient un quand il s'agit de comparer la répartition des résultats. En conséquence, les résultats obtenus à la formule 2 furent ajustés à ceux de la formule 1 (la procédure utilisée est définie dans le rapport du projet II). On doit donc avoir présent à l'esprit que lorsqu'on parle des éléments individuels, il s'agit des véritables réponses obtenues à ces éléments, tandis que lorsqu'on parle de n'importe quel résultat, il s'agit des résultats égalisés.

Le problème final est que les données ont été séparées de beaucoup de façons. Les résultats aux éléments individuels sont, dans la plupart des cas, différenciés seulement en fonction de l'année--12ème ou 13ème. Les données provenant des

tests-inventaires estimatifs figurent séparément dans le rapport selon l'année : 13ème année, 12ème année cours avancé, 12ème année cours général, 12ème année cours de base. La répartition des résultats figure dans le rapport pour cinq groupes différents, dont aucun ne correspond exactement aux quatre groupes que l'on vient de citer. Ces groupes sont :

- (a) les étudiants préparant le DSES (13ème année) et prévoyant de poursuivre un enseignement post-secondaire--universitaire dans la plupart des cas ;
- (b) les étudiants préparant le DSES et ayant d'autres projets pour l'année 1976-77 ;
- (c) les étudiants préparant le DES (12ème année) et continuant leurs études secondaires ;
- (d) les étudiants préparant le DES et prévoyant de poursuivre un enseignement post-secondaire--collège communautaire dans la plupart des cas ;
- (e) les étudiants préparant le DES et ayant d'autres projets pour l'année 1976-77.

On peut grossièrement assimiler le groupe (c) aux étudiants de 12ème année cours avancé, et les groupes combinés (d) et (e) aux étudiants de 12ème année cours général et cours de base, mais on ne peut présumer de l'exactitude de la concordance.

4. FORME DE L'ANALYSE

Tout d'abord, en raison de la négligeable concordance entre les données du projet II et celles du projet III, ce chapitre n'aura pas la présentation habituelle des résultats. Chacun des deux tests sera décrit et les résultats analysés en détail. Puis on

tentera d'établir autant que possible, une relation entre ces résultats et les objectifs définis dans le questionnaire sur les descriptions de cours qui apparaît dans le projet III.

5. LE TEST DE COMPREHENSION EN LECTURE ET DE CONNAISSANCE DE LA LANGUE

5.1 Description du test

Comme on l'a montré ci-dessus, ce test était divisé en deux formules approximativement parallèles. Chacune d'entre elles consistait en une partie A qui testait la compréhension de textes, et une partie B qui traitait de l'orthographe d'accord et du vocabulaire. Dans les deux cas, la partie A comprenait deux textes, suivi chacun de cinq à huit questions s'y rapportant. La partie B comprenait trois types différents d'éléments à choix multiple. Dans le premier type, décrit ici comme étant "éléments d'identification d'une phrase fautive", l'étudiant devait sélectionner un élément, le cas échéant, parmi quatre phrases données qui contenaient une erreur de grammaire ou de ponctuation. Le second type, appelé ici "éléments d'identification et de classification d'une erreur", présentait une phrase à l'étudiant qui devait déterminer si elle contenait une erreur grammaticale. Si oui, il devait la classer sous l'une des rubriques suivantes :

(A) Erreur à un groupe nominal : nom, adjectif, article (genre, nombre, accord).

(B) Erreur à un pronom.

(C) Erreur à un groupe verbal : auxiliaire, verbe (forme, temps, accord).

(D) Erreur à un adverbe, une préposition, une conjonction.

Le troisième groupe d'éléments appelé "phrases à compléter" contenait dans chaque cas une phrase incomplète. L'étudiant devait sélectionner parmi cinq propositions le mot ou groupe de mots qui convenait le mieux pour compléter la phrase.

Ainsi qu'il a été suggéré ci-dessus, le comité de sélection des tests ne fut pas entièrement satisfait du choix de ce test. Il était composé de sections extraites d'un test conçu à l'origine pour des étudiants qui se destinaient à faire des études universitaires et étaient en dernière année du secondaire. Le comité réalisa qu'il aurait pu être quelque peu difficile pour les étudiants de 12ème année, particulièrement pour ceux des cours généraux et des cours de base. Cependant il était, parmi tous les tests disponibles, celui qui convenait le mieux et le comité pensa qu'on pouvait en tirer des indications importantes en l'utilisant, même en considérant son degré élevé de difficulté.

5.2 Partie A : compréhension de textes

Le tableau 8.2 présente les réponses données par les professeurs du secondaire à un nombre de questions du test-inventaire estimatif qui traitent de la compréhension de textes et des éléments sur le test de langue destinés à mesurer les compétences des étudiants en la matière. Presque tous les professeurs, à la fois ceux de 13ème année et de tous les niveaux de 12ème année ont été d'accord sur le fait qu'un test de compréhension en lecture détermine une composante importante de la connaissance de la langue. On leur avait également demandé de répondre à des questions relatives aux trois composantes particulières de compétence en lecture que le test voulait mesurer.

La plupart des professeurs de tous les niveaux ont pensé qu'au moins trois-quarts de leurs étudiants devraient pouvoir lire de façon à avoir une compréhension littérale avec une proportion plus forte pour la 13ème année (19 des 20 professeurs) et plus

faible pour les professeurs de 12ème année cours de base. Pour tous les cours, excepté le cours de base 12ème année, au moins 70% des professeurs ont précisé qu'ils accordaient à cette compétence une attention marquée ou très marquée dans leur enseignement de la lecture. Les prévisions furent seulement légèrement inférieures dans le domaine de la lecture pour identifier l'idée principale ou les objectifs de l'auteur (excepté pour la 13ème année où les professeurs s'attendaient au moins à ce que trois-quarts de leurs étudiants aient cette faculté). On a mis un très fort accent sur cette compétence à tous les niveaux, excepté en 12ème année cours de base. Les prévisions furent également légèrement inférieures pour les étudiants capables de tirer des conclusions ou de dégager des implications et on a mis sur cette compétence un accent quelque peu moins important que sur les autres, à l'exception de la 13ème année.

On avait également demandé aux professeurs d'évaluer sur une échelle allant de "trop facile" à "trop difficile", chacun des quatre passages proposés et d'évaluer de la même façon les quatre passages dans leur ensemble. Les réponses détaillées sont indiquées dans le tableau 8.2. Les professeurs de 13ème année se sont mis d'accord pour trouver que, d'une façon générale, les passages étaient "bien choisis" par rapport à un niveau de difficulté. Les professeurs de 12ème année les ont considérés quelque peu plus difficiles avec une proportion les décrivant comme "difficiles" ou "trop difficiles", croissant du niveau avancé au niveau de base, en passant par le niveau général.

Le tableau 8.3 montre, pour chacun des quatre passages, l'évaluation moyenne des professeurs à chaque niveau (1 = "trop facile", 2 = "facile" ...) et pour chaque élément du test relatif au passage, le pourcentage d'étudiants de 12ème et de 13ème année qui ont fourni la réponse correcte.

Il n'est pas surprenant que les professeurs de 13ème année ont considéré trois des passages comme étant plus faciles, que ne l'ont pensé tous les groupes de professeurs de 12ème année. Un passage fut coté légèrement plus difficile par les professeurs de

13ème année que par les professeurs de 12ème année niveau avancé. Les professeurs de 13ème année ont coté trois passages à 3 ou proche de 3 ("bien choisi") et un comme légèrement plus difficile. A l'autre extrémité, les professeurs de 12ème année cours de base ont placé trois des quatre passages au niveau 4 ("difficile") et n'en ont déterminé qu'un comme "bien choisi".

Pas surprenant non plus que les résultats des étudiants de 13ème année sur tous les passages soient meilleurs que ceux des étudiants de 12ème année. En fait, il n'y a pas eu d'éléments de compréhension en lecture que les étudiants de 12ème année aient réussis aussi bien que ceux de 13ème année, bien que pour quelques-uns le taux de succès ait été très proche.

Ce qui est un peu surprenant est que trois des quatre groupes de professeurs ont considéré le quatrième passage comme étant le plus facile, et cependant, à tous les niveaux il s'est avéré être le plus difficile. L'explication plausible de ces faibles résultats relativement à ce passage est que celui-ci est écrit sur un ton particulièrement ironique et les étudiants n'ont pas perçu cette ironie. Les réponses données à un certain nombre d'éléments ont indiqué que les étudiants ont interprété le texte littéralement.

Si on examine les éléments en tant qu'expression concrète des trois compétences particulières en lecture mentionnées plus haut, on voit que 11 des 25 éléments mesurent d'abord le faculté de lire pour comprendre le texte de façon littérale. Le pourcentage moyen de réponses correctes pour ces 11 éléments était de 65 pour la 13ème année et de 53 pour le 12ème année. Les étudiants de 13ème année tombent en dessous de 50% pour un seul de ces éléments. Ceux de 12ème année en ajoutaient un second--un de ceux relatifs au quatrième passage (ironique). En raison de la difficulté du test, ceci semble être un résultat acceptable pour l'ensemble des niveaux d'étude et justifie les bonnes prévisions des professeurs.

Quatre éléments de test demandaient à l'étudiant d'identifier l'idée principale ou l'objectif d'un auteur. L'un d'entre eux fut celui qui obtint les meilleures réponses de tous les éléments de compréhension en lecture, avec 95% des réponses correctes pour les étudiants de 13ème année et 82% pour les étudiants de 12ème année. On tombe très proche de la moyenne avec 52% et 45% de succès respectivement pour les 13ème et 12ème année. Les deux autres ont trait au quatrième passage et les taux de succès pour les deux niveaux d'étude ont été très bas. Ainsi que nous l'avons mentionné ci-dessus, les réponses données à ces deux éléments indiquaient que beaucoup d'étudiants ont interprété le passage littéralement sans en percevoir le ton ironique. Les résultats apparaîtraient ici quelque peu inférieurs à ce à quoi s'attendaient les professeurs et indiqueraient un manque d'appréciation de la subtilité du ton. C'est un domaine vers lequel les professeurs devraient porter leur attention.

Dix éléments demandaient aux étudiants de tirer des conclusions ou de dégager des implications d'un passage donné. Les taux moyens de succès à ces éléments (57% pour la 13ème année et 46% pour la 12ème année) étaient légèrement inférieurs que pour la compréhension littérale d'un texte et légèrement supérieurs que pour l'identification de l'idée principale ou l'objectif d'un auteur. Le fait que les résultats en ce domaine soient quelque peu plus pauvres que pour la compréhension littérale n'est pas surprenant, étant donné les faibles prévisions des professeurs et le faible accent mis sur cette compétence.

Le rapport du projet II contient des comparaisons de résultats sur cette partie du test pour les cinq groupes d'étudiants décrits plus haut (selon le niveau d'études et les orientations pédagogiques). Comme on pouvait l'espérer les résultats étaient plus élevés pour les étudiants préparant le DSES et prévoyant de poursuivre un enseignement post-secondaire. Les autres étudiants préparant le DSES et ceux du DES restant dans un établissement secondaire ont eu des résultats légèrement inférieurs et grossièrement comparables. Les étudiants préparant le DES et prévoyant de poursuivre un enseignement post-secondaire

se sont placés un peu plus bas sur l'échelle et les plus faibles résultats ont été obtenus par des étudiants préparant le DES mais ayant autres perspectives.

5.3 Partie B : mécanismes d'écriture

Le tableau 8.4 montre les réponses des professeurs aux questions du test-inventaire estimatif relatives à la partie B du test de langue et des aptitudes que ce test cherchait à mesurer. Tous les types d'éléments testant l'orthographe d'accord furent jugés "tout à fait appropriés" par la plupart des professeurs de 13^{ème} année. Les appréciations des professeurs de 12^{ème} année cours avancés concordaient. Les professeurs de 12^{ème} année cours généraux ont été plus hésitants quant aux éléments du second type, les jugeant quelque peu difficiles quoique, par ailleurs, appropriés. La plupart des professeurs de 12^{ème} année cours de base les ont trouvés trop difficiles. Les éléments testant le vocabulaire furent considérés de la même façon que le second type d'éléments d'orthographe. Ils convenaient aux étudiants de 13^{ème} année et de 12^{ème} année cours avancés mais étaient trop difficiles pour les autres étudiants de 12^{ème} année.

Les professeurs de tous les niveaux d'études ont pensé que les éléments fournissaient un bon équilibre en testant des secteurs importants de la connaissance de la langue. Au moins 65% des professeurs de tout niveau ont donné aux aptitudes testées une attention marquée ou très marquée dans leur enseignement. La proportion d'étudiants ayant--selon les prévisions des professeurs--des compétences au début de leur dernière année du secondaire tombe brusquement à mesure que le niveau d'étude baisse. Bien que 15 de 20 professeurs de 13^{ème} année pensaient qu'au moins trois-quarts de leurs étudiants devraient avoir ces compétences, seulement 1 des 5 professeurs de 12^{ème} année cours de base avait fait cette prévision. Parmi 19 professeurs de 13^{ème} année qui ont répondu, 17 pensaient que trois-quarts de leurs étudiants devraient être compétents en ce domaine à la fin de

cette année-là. Au niveau de 12ème année cours de base, 3 des 5 professeurs ont donné la même réponse.

Le tableau 8.5 contient une liste des éléments du test ayant trait à l'orthographe (groupes A et B) et au vocabulaire (groupe C). On y verra, pour chaque élément, le pourcentage d'étudiants de 12ème et de 13ème année ayant fourni les réponses correctes. Comme il est indiqué dans le tableau, deux des éléments n'ont pas été évalués parce que l'équipe du projet II a réalisé que ces éléments offraient deux possibilités de réponses correctes.

En premier lieu, si nous examinons les résultats des tests sur les 23 éléments évalués dans les deux premiers groupes, il apparaît clairement que les éléments présentaient pour les étudiants un grand éventail de difficulté. Les taux de succès vont de 9% à 70% pour les étudiants de 13ème année et de 14% à 63% pour les étudiants de 12ème année. On a examiné chaque élément dont les résultats se sont révélés relativement faibles (en-dessous de 40% pour le niveau de 13ème année) de façon à déterminer, si possible, la cause particulière de la difficulté de l'élément (bien qu'il soit impossible d'établir avec certitude les raisons pour lesquelles les étudiants sont attirés par une réponse fausse ou évitent la réponse correcte, il est souvent possible de faire des suppositions hautement probables).

Dans l'élément I-13 (formule 1, élément 13) beaucoup d'étudiants de tous les niveaux ont considéré comme incorrecte une phrase contenant un verbe familier utilisé dans un sens un peu moins familier. L'acception la plus familière ne convenait pas au contexte. Il apparaît que les étudiants ont pensé que le verbe était utilisé de façon incorrecte.

L'élément I-16 demandait que l'étudiant reconnaisse l'usage incorrect dans une expression idiomatique. L'expression proposée contenait le mot "au" qui donnait à l'expression une signification différente de celle qui convenait au contexte et était employé à la place du mot "à". Les étudiants les plus

faibles semblent avoir choisi comme erronée une phrase contenant un sujet composé dont chacune des parties du sujet était au singulier, suivi d'un verbe au pluriel. Les différentes parties du sujet étaient presque synonymes et les étudiants ont pu penser que le verbe devait être au singulier. Les autres erreurs étaient réparties à peu près uniformément sur les questions plus ambiguës.

L'élément I-18 fut le plus mal fait relativement à l'orthographe d'accord. C'est le seul de ce groupe où les étudiants de 12ème année aient obtenu de meilleurs résultats que les étudiants de 13ème année. Il demandait aux étudiants de réaliser qu'en pluralisant un nom composé formé d'un adjectif et d'un nom, on doit aussi pluraliser l'adjectif (exemple : monsieur--messieurs). Les erreurs étaient par ailleurs diversement réparties sur les autres réponses.

Dans l'élément I-19, les étudiants de tous les niveaux ont été nombreux à relever en particulier deux fausses réponses. Toutes deux comportaient des mots correctement orthographiés qui auraient pu aisément être considérés par un étudiant faible ou négligent comme ne s'accordant pas en genre avec les noms qu'ils modifiaient apparemment.

Dans l'élément I-20, les étudiants ont semblé avoir trouvé l'erreur mais y ont assigné la mauvaise classification. L'erreur était située dans un groupe de mots contenant un verbe et un adverbe; les étudiants l'ont classée comme verbale plutôt que comme adverbiale.

Dans l'élément I-23, les étudiants ont laissé passer un article dont le genre était incorrect. En place, ils ont sélectionné comme erroné un verbe singulier suivant un sujet de la forme "chacun des ...".

L'élément II-15 contenait une petite erreur typographique qui explique probablement partiellement les faibles résultats obtenus. L'erreur concernait une lettre majuscule et beaucoup

d'étudiants relativement faibles ont choisi cette phrase en question et ont laissé au contraire passer une erreur beaucoup plus sérieuse dans une forme verbale qui se trouvait dans une autre phrase.

L'élément II-17 présentait aux étudiants une tâche un peu compliquée de discernement. Deux des réponses incorrectes contenaient des noms collectifs, l'un dûment suivi d'un verbe au singulier et l'autre dûment suivi d'un verbe au pluriel. Les étudiants ont choisi les deux en grande nombre et ont manqué une erreur plus nette dans un accord nom-verbe qui était le choix correct.

Un grand nombre d'étudiants n'ont pas vu d'erreur dans les éléments II-23, II-24 et II-26. En fait, le premier contenait une erreur de genre, le second une forme incorrecte de l'impératif d'un verbe irrégulier et le troisième un ordre fautif d'un verbe à l'impératif, son complément d'objet direct et son complément d'objet indirect.

Quatre des erreurs faites par beaucoup d'étudiants étaient de nature telle que la forme incorrecte donnée dans le test était phonétiquement équivalente à la forme correcte.

En récapitulant, des onze éléments ayant eu de très faibles résultats, trois contenaient un accord sujet-verbe avec une structure de sujet compliquée; deux mettaient en évidence des verbes à l'impératif; un traitait de la pluralisation d'un nom composé; deux contenaient des erreurs de genre; un contenait un mot correct en genre mais considéré comme faux par les étudiants; un utilisait un mot usuel de façon non commune que les étudiants ont considéré comme faux; un demandait aux étudiants de faire une distinction entre deux expressions idiomatiques ne différant respectivement que par l'usage de "à" et "au". Dans le dernier élément, les étudiants ont reconnu l'erreur mais ont été incapables de la classer correctement.

Les éléments inclus dans le groupe C au tableau 8.5 étaient ceux demandant aux étudiants de choisir les mots corrects pour compléter une phrase donnée; ils furent classés sous la rubrique "vocabulaire". Là encore, les étudiants de 13^{ème} année ont logiquement fait mieux que les étudiants de 12^{ème} année et ce avec une plus grande marge que pour les éléments d'orthographe d'accord. A propos de ces derniers éléments, le taux de succès de 13^{ème} année était en moyenne supérieur de 7% au taux de 12^{ème} année; le chiffre est ici de 13%.

Pour huit de ces éléments, le taux de succès pour les étudiants de 13^{ème} année est tombé en-dessous de 40% avec, en correspondance, des chiffres bas pour la 12^{ème} année. Dans l'un de ces cas, la difficulté semble résulter de la confusion avec l'anglais. Beaucoup d'étudiants ont choisi chacune des deux réponses incorrectes; dans chaque cas, le mot apparenté anglais avait le sens requis par le contexte. Dans les sept autres éléments, les réponses données font apparaître que les étudiants n'étaient pas familiers avec le vocabulaire requis, ou étaient incapables de l'utiliser de façon appropriée. La répartition des réponses ressemble à celle qui aurait pu résulter d'une simple estimation due au hasard.

On relève quelques faiblesses à la fois dans l'orthographe d'accord et dans le vocabulaire de cette section du test. Les résultats de 12^{ème} année en orthographe requièrent moins de considération parce que, répétons-le, le test n'était pas originellement conçu pour les étudiants de ce niveau. Cependant, la différence notoire entre les résultats des étudiants de 12^{ème} année et ceux de 13^{ème} année dans la section du vocabulaire laisse penser que beaucoup d'étudiants de 12^{ème} année dans les cours généraux ou les cours de base sont faibles en ce domaine. Les résultats de 13^{ème} année ne donnent pas lieu de s'inquiéter outre mesure mais indiquent que certains secteurs justifieraient une amélioration.

6. LE TEST DE COMPOSITION ECRITE

6.1 Description du test

Le test de composition écrite fut une solution plus heureuse aux problèmes de la diversité des étudiants et des cours que ne le fut le test de langue. L'étudiant devait écrire un essai sur l'un des huit sujets proposés. (Ceci fut précisément le même travail que celui présenté aux étudiants testés en anglais dans les écoles anglophones, avec le même choix de sujets.) Ces essais furent corrigés, dans la plupart des cas, par trois personnes à qui on avait demandé de donner une note générale à chaque essai en utilisant une échelle allant de 1 à 10. (Les résultats pour chacun des cinq groupes résultant d'une répartition selon le niveau d'études et les orientations pédagogiques figurent dans le rapport du projet II.)

6.2 Réponses aux questions du test-inventaire estimatif

Le tableau 8.6 présente les réponses des professeurs aux questions du test-inventaire estimatif relatives au test de composition écrite, l'accent qu'ils mettent sur cette aptitude et leurs prévisions concernant les résultats de leurs étudiants. La majorité des professeurs, excepté ceux de 12ème année cours de base, mettent une grande importance sur un échantillon de composition d'étudiants dans l'évaluation de la compétence linguistique. La plupart pensent que cet échantillon de compositions devrait être complété par un test de langue à choix multiple, bien que beaucoup donneraient plus de poids à la composition écrite.

La proportion des professeurs qui pensaient que leurs étudiants pourraient présenter une composition acceptable sur les sujets proposés décroît à mesure que le niveau d'études baisse. Des 20 professeurs de 13ème année, 18 pensaient que plus des trois-quarts de leurs étudiants pourraient s'acquitter de ce travail; des 5 professeurs de 12ème année cours de base, deux

seulement avaient fait cette prévision. Seuls les professeurs de 13^{ème} année et de 12^{ème} année cours avancé étaient convaincus que le niveau de difficulté de ce test était raisonnable; environ la moitié des professeurs de 12^{ème} année cours généraux et cours de base étaient d'accord. Pour la plupart, les professeurs pensaient qu'on ne désavantageait pas les étudiants en ne leur permettant pas de choisir leur mode d'expression; ils ont approuvé le choix des sujets, bien que beaucoup ont pensé qu'il y avait d'autres modes au moins aussi importants, si ce n'est plus, que celui testé. Un accent plus grand était mis sur ce type de composition dans les cours de 13^{ème} année et dans ceux de 12^{ème} année niveau avancé que dans les cours des autres niveaux de 12^{ème} année.

On avait demandé aux professeurs de classer par ordre d'importance cinq critères qui pouvaient être utilisés pour évaluer une telle composition écrite. Les professeurs de 13^{ème} année ont mis plus d'accent sur le style, moins sur le choix des mots. Pour ceux de 12^{ème} année niveau avancé, la structure générale était plus importante avec, de nouveau, le choix des mots moins important. Les professeurs de 12^{ème} année cours généraux ont placé la présentation logique des arguments en premier, le choix des mots en dernier. Le plus important pour les professeurs de 12^{ème} année cours de base était la technique.

6.3 Analyse des compositions écrites

Le rapport du projet II contient en appendice B une étude détaillée de quelques-uns de ces essais (Mougeon, Bélanger, Canale). Cette section traite d'un échantillon de 50 rédactions choisies au hasard qui ont été notées par neuf correcteurs au lieu de trois (pour assurer une plus grande fiabilité des résultats), et d'un sous-groupe de 16 rédactions plus en détail. Il reste peu à dire ici sur ce sujet de la compétence des étudiants à composer et nous rapporterons seulement brièvement quelques-unes des conclusions trouvées dans cette section.

Pour l'échantillon des 50 rédactions, les indices d'erreurs ont été calculés en divisant le nombre total d'erreurs commis par un étudiant par le nombre total de mots écrits par l'étudiant. En moyenne, ces indices étaient beaucoup plus bas pour les étudiants de 13ème année et pour ceux de 12ème année continuant vers la 13ème année que pour les autres étudiants de 12ème année. Cependant, l'index d'erreurs moyen pour les étudiants de 13ème année continuant vers l'université était de 0.10--ou une erreur tous les dix mots--et plusieurs de ces étudiants avaient des indices d'erreurs particulièrement élevé (0.15 ou plus). L'index d'erreurs moyen pour les étudiants de 12ème année restant au secondaire était à peu près le même : 0.09, mais on n'a pas trouvé là d'indices très élevés. Les indices les plus hauts apparaissent avec les étudiants de 12ème année entrant sur le marché du travail : ces étudiants ont en moyenne une erreur tous les cinq mots.

On a fait une comparaison entre les indices d'erreurs de ces 50 rédactions et les notes attribuées à ces rédactions par les correcteurs. La relation entre ces deux éléments était très proche.

Seize de ces rédactions ont été analysées très en détail de façon à déterminer les types d'erreurs et leur relative fréquence. Les erreurs d'usage apparaissent plus fréquemment avec les prépositions puis avec les verbes. Les erreurs d'accord étaient rares dans l'utilisation de l'article, plus fréquentes avec les adjectifs et les verbes où la langue parlée ne fait pas toujours la différence entre les formes variées d'un mot.

Une grande proportion d'erreurs d'orthographe étaient attribuables à ce manque de distinction entre deux mots dans le langage parlé. Des 174 erreurs d'orthographe trouvées dans les 16 rédactions, 77 étaient de ce type. D'autre part, les erreurs d'usage étaient attribuables dans beaucoup de cas à l'influence de l'anglais sur l'intelligence de sa propre langue (par exemple, dans la façon d'utiliser les prépositions).

Les questions qui ne ressortaient pas directement de la correction de la grammaire et de l'orthographe n'ont pas été examinées de façon approfondie. On a noté qu'en matière de style, certains élèves ont utilisé une quantité non négligeable de structures caractéristiques de la langue parlée informelle, et ceci fut attribué--au moins partiellement--au fait que de par leur statut minoritaire, les jeunes Franco-ontariens sont assez peu exposés aux variétés de français formel. Sur la question de la présentation logique des idées, les correcteurs ont fait ce commentaire : "signalons que seule une infime minorité des étudiants a fait preuve d'une telle capacité. En fait, la plupart du temps, nous avons constaté que les idées sont mal enchaînées."

7. SOMMAIRE

A ce point, nous revenons à la liste des objectifs incluse dans le questionnaire sur les descriptions de cours, dans le but de classer les données des tests du projet II sous les rubriques de cette liste.

Le tableau 8.7 donne, pour chacun des objectifs pour lesquels des données du projet II sont disponibles, le niveau de sortie moyen des étudiants indiqué par les professeurs ayant rempli le questionnaire. L'impression générale qui ressort de ce tableau est que, selon leurs professeurs, les étudiants de 13^{ème} année ont des résultats approchant le niveau "très acceptable"; les étudiants de 12^{ème} année cours avancés ont également à peu près ce niveau. Les professeurs de 12^{ème} année cours généraux évaluent la compétence de leurs étudiants comme étant légèrement inférieure, mais encore bien au-dessus du niveau acceptable. Nous tenterons de comparer ces prévisions avec les résultats réels des étudiants par rapport à ces objectifs.

7.1 Orthographe

Comme nous l'avons mentionné ci-dessus, l'orthographe est testée par le test à choix multiple seulement dans le contexte de l'orthographe d'accord. Nous considérerons ces résultats sous la rubrique "grammaire". Dans les rédactions du test de composition écrite, les erreurs d'orthographe ont été fréquentes. Un grand nombre de ces erreurs pourraient être attribuées au manque de distinction dans le français parlé entre les formes correctes et erronées.

7.2 Vocabulaire

C'est dans le domaine du vocabulaire qu'on enregistre les plus grandes différences entre les résultats de 12ème et de 13ème année dans le test de langue. Les résultats de 13ème année ont indiqué quelques points faibles à ce sujet. Sur la base de comparaisons des résultats (voir le rapport du projet II), on peut conclure que les étudiants de 12ème année cours avancé ont des résultats assez proches du niveau de 13ème année et que la grande différence entre les résultats globaux pour les 12ème et 13ème année doit être attribuée aux manques sérieux de vocabulaire chez les étudiants de 12ème année inscrits aux cours généraux et aux cours de base. Ceci est corroboré par les auteurs de l'Appendice B du rapport du projet II qui font ce commentaire : "(certains élèves) ont un vocabulaire très approximatif".

7.3 Grammaire

Les résultats des étudiants sur la partie du test de langue relative à la grammaire (particulièrement l'orthographe d'accord) ont montré quelques faiblesses. On trouve la confirmation de l'hypothèse mentionnée ci-dessus sous le titre "orthographe" que les étudiants ont de la difficulté avec les mots qui s'orthographient différemment mais qui se prononcent de la même façon; quatre des éléments pour lesquels on a enregistré de

faibles résultats contenaient des problèmes de ce type. D'autre part, le rapport sur les compositions écrites a montré que les erreurs d'accord d'un article avec le nom qui lui est rattaché arrivent rarement. Cependant ce genre d'erreurs apparaît dans un des éléments à choix-multiple ayant obtenu de faibles résultats. Dans un autre, on trouve une erreur d'accord entre le nom et l'adjectif démonstratif (où s'appliqueraient quelques-uns des arguments identiques qui ont été utilisés pour expliquer la non-fréquence relative des erreurs d'articles). L'accord sujet-verbe a semblé être un problème, particulièrement là où la forme du sujet était quelque peu compliquée.

Les informations provenant du travail écrit des étudiants montrent que beaucoup d'erreurs d'usage peuvent être attribuées à la contamination par l'anglais. Ceci s'applique en particulier au choix erroné des prépositions, à leur omission ou à leur introduction dans des conditions inappropriées; ceci s'applique également à l'usage incorrect des verbes.

7.4 Stylistique

On a porté peu d'attention à la stylistique en tant que question indépendante. Les impressions subjectives des auteurs de l'Appendice B du rapport du projet II montrent que les étudiants sont peu habiles à faire une distinction entre les niveaux de style et tendent à utiliser beaucoup d'expressions et de structures acceptables dans le français informel, dans des situations où un usage plus formel serait recommandé. Ils attribuent largement ceci au manque d'exposition à un niveau formel de langue qui est le résultat inévitable du statut minoritaire des Franco-ontariens.

7.5 Travaux écrits : organisation et rédaction

Les commentaires faits dans le rapport du projet II au sujet des aptitudes des étudiants à présenter leurs idées d'une façon

logique étaient plus qu'un peu pessimiste. Il apparaît qu'il s'agit d'un secteur nécessitant une grande amélioration, de façon à ce que plus de "l'infime minorité" des étudiants mentionnée dans ce rapport puissent quitter l'école secondaire en sachant présenter un argument de façon organisée et logiquement cohérente.

7.6 Analyse et commentaires des textes

Cet objectif est inclus dans la liste des sujets sous la rubrique "travaux écrits". Il n'a pas été examiné dans ce contexte, mais traité dans la section du test de langue intitulée "compréhension de textes". Les résultats se sont situés à un niveau acceptable, excepté dans le cas d'un passage où beaucoup d'étudiants n'ont pas perçu le ton ironique. On recommande de mettre un peu l'accent sur l'aptitude à apprécier de telles subtilités de ton dans le langage écrit.

TABLEAU 8.1

OBJECTIFS INCLUS DANS LE QUESTIONNAIRE SUR LES DESCRIPTIONS DE COURS (PROJET NO. III)

A. Langue:

1. Etude des éléments fondamentaux de la langue:

- | | |
|--|---|
| 1.1 orthographe | * |
| 1.2 vocabulaire | * |
| 1.3 grammaire | * |
| 1.4 stylistique (structures de phrases, phraséologie etc.) | * |

2. Travaux écrits:

- | | |
|--------------------------------------|---|
| 2.1 compréhension du sujet | * |
| 2.2 organisation | * |
| 2.3 rédaction | * |
| 2.4 analyse et commentaire de textes | * |

3. Expression orale:

- | |
|--------------------------|
| 3.1 prononciation |
| 3.2 lecture expressive |
| 3.3 expression spontanée |
| 3.4 exposés |

TABLEAU 8.1 (suite)

B. Linguistique:

1. Principes généraux

2. Domaines:

2.1 phonétique

2.2 phonologie

2.3 morphologie

2.4 syntaxe

2.5 sémantique

2.6 lexicologie

C. Littérature:

1. Etude de textes:

1.1 poèmes

1.2 romans

1.3 théâtre

1.4 essais

2. Explication de textes:

2.1 orale

2.2 écrite

*

TABLEAU 8.1 (suite)

3. Composition:

- 3.1 analyse littéraire
- 3.2 commentaire de textes
- 3.3 dissertations
- 3.4 rédactions diverses

D. Culture:

1. Connaissance de l'héritage littéraire et culturel:

- 1.1 poèmes
- 1.2 romans
- 1.3 contes
- 1.4 théâtre
- 1.5 films
- 1.6 chansons

TABLEAU 8.2

REPONSES DES ENSEIGNANTS AUX QUESTIONS CONCERNANT
LES TEXTES RELATIFS A LA COMPREHENSION EN LECTURE

	<u>13ème</u> <u>année</u>	<u>niveau</u> <u>avancé</u>	<u>12ème année</u> <u>niveau</u> <u>général</u>	<u>niveau</u> <u>de base</u>
Considérez-vous qu'un test de compréhension en lecture-- pas nécessairement celui utilisé ici--détermine une composante importante de la connaissance de la langue (à votre niveau)?				
Oui	18	15	15	4
Non	1	0	0	0

Si l'on considère les deux
formules du test, il y a
quatre textes relatifs à
la compréhension en lecture.
Quelle est votre opinion
générale quant au niveau
de difficulté de ces quatre
textes?

trop facile	0	0	0	0
facile	2	0	0	0
bien choisi	13	8	5	0
difficile	4	6	4	3
trop difficile	0	0	5	1

...du premier texte?

trop facile	2	0	0	0
facile	4	2	2	0
bien choisi	9	9	6	0
difficile	4	4	4	4
trop difficile	1	0	3	1

...du deuxième texte?

trop facile	0	0	0	0
facile	2	0	0	0
bien choisi	9	9	5	1
difficile	7	6	5	2
trop difficile	2	0	5	2

TABLEAU 8.2 (suite)

	13ème année	12ème année		
		niveau avancé	niveau général	niveau de base
...du troisième texte?				
trop facile	1	0	0	0
facile	2	2	1	1
bien choisi	13	7	4	1
difficile	4	4	7	1
trop difficile	0	2	3	2
...du quatrième texte?				
trop facile	0	0	0	0
facile	7	2	2	1
bien choisi	11	7	7	2
difficile	2	5	3	2
trop difficile	0	1	3	0
Combien parmi les étudiants commençant les cours à (votre) niveau devraient pouvoir comprendre littéralement un texte?				
100%	11	6	5	1
76%-99%	8	7	6	2
51%-75%	1	2	4	1
26%-50%	0	0	0	1
1%-25%	0	0	0	0
aucun	0	0	0	0
Quelle attention accordez-vous en lecture à la compréhension littérale du texte?				
attention très marquée	10	7	5	0
attention marquée	4	5	8	2
attention modérée	6	3	2	3
attention particulière à un étudiant plus faible	0	0	0	0
attention nulle	0	0	0	0

TABLEAU 8.2 (suite)

	13ème année	12ème année		
		niveau avancé	niveau général	niveau de base
Combien parmi les étudiants commençant les cours à (votre) niveau devraient pouvoir identifier l'idée principale ou les objectifs de l'auteur?				
100%	10	5	3	2
76%-99%	10	7	7	0
51%-75%	0	3	5	3
26%-50%	0	0	0	0
1%-15%	0	0	0	0
aucun	0	0	0	0

Quelle attention accordez-vous
en lecture à l'identification
de l'idée principale ou des
objectifs de l'auteur?

attention très marquée	11	4	3	1
attention marquée	6	7	8	0
attention modérée	3	4	4	4
attention particulière à un étudiant plus faible	0	0	0	0
attention nulle	0	0	0	0

Combien parmi les étudiants
commençant les cours
à (votre) niveau devraient
pouvoir tirer des conclusions
ou dégager des implications?

100%	7	3	1	0
76%-99%	10	5	8	3
51%-75%	3	4	6	2
16%-50%	0	3	0	0
1%-25%	0	0	0	0
aucun	0	0	0	0

TABLEAU 8.2 (suite)

		12ème année		
	13ème année	niveau avancé	niveau général	niveau de base
Quelle attention accordez-vous en lecture au pouvoir de tirer des conclusions ou de dégager des implications?				
attention très marquée	10	2	3	0
attention marquée	8	7	6	2
attention modérée	2	6	6	3
attention particulière à un étudiant plus faible	0	0	0	0
attention nulle	0	0	0	0

TABLEAU 8.3

RESULTATS DE LA PARTIE "A" DU
TEST DE COMPREHENSION EN LECTURE ET DE CONNAISSANCE DE LA LANGUE (FRANCAIS)

Texte	Moyenne des évaluations du niveau de difficulté*			Elément(s)	% de réponses correctes	
	13ème	12ème av.	12ème gén.		13ème	12ème
1	2.9	3.1	3.5	4.2	I-1	72
					I-2	67
					I-3	68
					I-4	70
					I-5	74
					Moyenne	70
2	3.5	3.4	4.0	4.2	I-6	75
					I-7	67
					I-8	26
					I-9	58
					I-10	52
					I-11	38
Moyenne					53	
3	3.0	3.4	3.8	3.8	II-1	76
					II-2	78
					II-3	83
					II-4	67
					II-5	83
					II-6	69
Moyenne					72	
Moyenne					72	

TABLEAU 8.3 (suite)

Texte	Moyenne des évaluations du niveau de difficulté*			Elément(s)	% de réponses correctes	
	13ème	12ème av.	12ème gén.	12ème de base	13ème	12ème
4	2.8	3.3	3.5	3.2	68	58
				II-9	51	33
				II-10	42	34
				II-11	28	20
				II-12	36	25
				II-13	30	19
				II-14	43	32
				Moyenne		

- *1 - trop facile
2 - facile
3 - bien choisi
4 - difficile
5 - trop difficile

TABLEAU 8.4

REPONSES DES ENSEIGNANTS AUX QUESTIONS CONCERNANT
LES ELEMENTS DU TEST RELATIFS A LA CONNAISSANCE DE LA LANGUE

	13ème année	12ème année		
		niveau avancé	niveau général	niveau de base
Quelle est votre évaluation en ce qui concerne les éléments d' <u>identification d'une</u> <u>phrase fautive?</u>				
appropriés mais trop faciles	2	3	0	0
tout à fait appropriés	16	9	11	2
appropriés mais trop difficiles	1	1	3	3
inappropriés pour des raisons autres que la difficulté	1	2	1	0
Quelle est votre évaluation en ce qui concerne les éléments d' <u>identification</u> <u>et de classification d'une</u> <u>erreur?</u>				
appropriés mais trop faciles	0	2	0	0
tout à fait appropriés	15	9	8	1
appropriés mais trop difficiles	3	1	6	3
inappropriés pour des raisons autres que la difficulté	2	3	1	1
Quelle est votre évaluation en ce qui concerne les éléments de <u>phrases à compléter?</u>				
appropriés mais trop faciles	0	2	0	0
tout à fait appropriés	16	11	9	2
appropriés mais trop difficiles	2	0	6	3
inappropriés pour des raisons autres que la difficulté	2	2	0	0

TABLEAU 8.4 (suite)

	13ème année	12ème année		
		niveau avancé	niveau général	niveau de base
Quelle est votre évaluation d'ensemble quant au niveau de difficulté de la partie des tests ayant trait à la connaissance de la langue?				
trop facile	0	1	0	0
facile	4	1	1	1
bien choisi	13	9	7	1
difficile	3	4	4	3
trop difficile	0	0	3	0
Quelle est votre évaluation des éléments des tests relatifs à la connaissance de la langue du point de vue de l'accent mis sur l'usage, le style, la grammaire, la structure des phrases, les expressions idiomatiques?				
Ils fournissent un bon équilibre en testant des secteurs importants de la connaissance de la langue.	16	11	13	4
Bien que testant un certain nombre de secteurs importants, l'accent n'est pas mis de façon équilibrée sur les différents éléments.	1	3	0	0
Des secteurs importants de la connaissance de la langue sont omis ou testés trop superficiellement.	2	1	1	1

TABLEAU 8.4 (suite)

	13ème année	12ème année		
		niveau avancé	niveau général	niveau de base
Combien d'élèves devraient avoir les compétences estimées dans les tests de connaissance de la langue en <u>entrant</u> dans (votre) cours?				
100%	6	2	2	1
76%-99%	9	8	3	0
51%-75%	5	4	7	2
26%-50%	0	1	3	2
1%-25%	0	0	0	0
aucun	0	0	0	0

Quelle attention accordez-vous
au développement des
compétences requises pour les
tests de connaissance de la
langue?

attention nulle--les compétences requises sont sans intérêt	0	0	0	0
attention très marquée	4	7	3	1
attention marquée	9	5	7	2
attention modérée	5	3	3	2
attention particulière à un étudiant plus faible	2	0	1	0
attention nulle--les compétences requises sont trop avancés	0	0	1	0

Combien d'étudiants qui
terminent avec succès (votre)
cours de français devraient
avoir les compétences évaluées
par les tests de connaissance
de la langue?

100%	10	7	4	1
76%-99%	7	6	5	2
51%-75%	2	2	3	1
26%-50%	0	0	3	1
1%-25%	0	0	0	0
aucun	0	0	0	0

TABLEAU 8.5

POURCENTAGE D'ETUDIANTS CHOISSANT LA REPONSE JUSTE
A CHAQUE ELEMENT DU TEST RELATIF A LA CONNAISSANCE
DE LA LANGUE

<u>Elément du test</u>	<u>13ème année</u>	<u>12ème année</u>
A. Eléments d'identification d'une phrase fautive		
I-12	44	31
I-13	33	25
I-14	*	*
I-15	64	54
I-16	21	14
I-17	52	36
I-18	9	14
I-19	15	19
II-15	51	38
II-16	29	27
II-17	26	20
II-18	44	36
II-19	43	32
II-20	*	*
II-21	51	48
B. Eléments d'identification et de classification d'une erreur		
I-20	26	18
I-21	42	25
I-22	70	61
I-23	20	19
I-24	42	36
II-22	69	63
II-23	35	35
II-24	33	29
II-25	45	37
II-26	17	12

TABLEAU 8.5 (suite)

C. Eléments de phrases à compléter

I-25	67	45
I-26	60	44
I-27	19	10
I-28	34	36
I-29	17	7
I-30	32	19
I-31	21	12
I-32	30	25
I-33	50	47
I-34	30	16
I-35	20	14
II-27	48	26
II-28	66	50
II-29	74	63
II-30	57	46
II-31	76	57
II-32	48	26
II-33	54	34
II-34	41	26
II-35	53	31

*L'équipe du projet a jugé que deux des éléments des tests pouvaient offrir deux réponses correctes. Les réponses choisies à ces deux éléments ne sont incluses dans aucune analyse.

TABLEAU 8.6

REPONSES DES ENSEIGNANTS AUX QUESTIONS CONCERNANT
LE TEST DE COMPOSITION ECRITE

		12ème année		
	13ème année	niveau avancé	niveau général	niveau de base
Quelle importance accordez-vous à un exemple de la composition de l'étudiant dans une évaluation de la compétence au niveau de la langue?				
importance très marquée	8	8	2	1
importance marquée	10	5	8	1
importance modérée	1	1	4	3
importance nulle	0	0	1	0
Comment considérez-vous l'utilisation, à la fois d'un test de connaissance de la langue avec questions à choix multiple et un exemple de composition, dans l'évaluation de la compétence au niveau de la langue?				
la formule de questions à choix multiple est satisfaisante en elle-même	0	0	1	0
le recours aux deux formules est important	15	12	14	4
la formule d'un exemple de composition est satisfaisante en elle-même	4	1	0	0
aucune des deux formules n'est satisfaisante	0	1	0	1

TABLEAU 8.6 (suite)

	13ème année	12ème année		
		niveau avancé	niveau général	niveau de base
Si, par rapport à un ensemble, les résultats d'un étudiant obtenus au test de composition écrite étaient différents de ceux obtenus au test à choix multiple (du type de ceux que l'on fait passer dans cette étude), quels sont ceux que vous considéreriez les plus valables en tant que mesure de la compétence d'un étudiant au niveau de la connaissance de la langue?				
les résultats obtenus à la composition écrite	4	4	0	0
les résultats obtenus au test à choix multiple	1	1	1	2
des résultats qui combinerait les deux avec l'accent mis sur la composition écrite	9	4	5	2
des résultats qui combinerait les deux avec l'accent mis sur le test à choix multiple	0	0	4	0
des résultats qui combinerait les deux de façon équivalente	6	6	5	1
Combien d'étudiants devraient pouvoir écrire un essai correct--du type présenté ici--à son entrée dans les cours de français (à votre niveau)?				
100%	11	2	0	0
76%-99%	7	8	7	2
51%-75%	2	3	5	1
26%-50%	0	1	2	0
1%-25%	0	1	0	2
aucun	0	0	1	0

TABLEAU 8.6 (suite)

	12ème année			
	<u>13ème</u> <u>année</u>	<u>niveau</u> <u>avancé</u>	<u>niveau</u> <u>général</u>	<u>niveau</u> <u>de base</u>
Quelle attention accordez-vous, dans votre enseignement, au développement de la compétence des étudiants dans ce type de composition écrite?				
attention très marquée	7	4	1	0
attention marquée	7	8	4	2
attention modérée	6	3	9	3
attention nulle	0	0	1	0

Combien d'étudiants qui
terminent avec succès les
cours de français à (votre)
niveau devraient pouvoir
écrire un essai correct du
type de ceux proposés dans
ce test?

100%	12	4	1	0
76%-99%	6	10	6	2
51%-75%	2	1	5	1
26%-50%	0	0	1	2
1%-25%	0	0	1	0
aucun	0	0	1	0

Y a-t-il d'autres modes de
composition écrite qui, du
point de vue d'une culture
général, seraient aussi ou
plus importants que celui
utilisé ici?

Oui	6	9	11	3
Non	12	5	4	2

La composition écrite
était-elle située à un
niveau raisonnable de
difficulté pour des
étudiants suivant les
cours de (votre) niveau?

Oui	19	15	8	2
Non	1	0	7	3

TABLEAU 8.6 (suite)

	13ème année	12ème année		
		niveau avancé	niveau général	niveau de base
Le fait de n'avoir proposé qu'un seul type de composition écrite était-il de nature à défavoriser les étudiants?				
Oui	4	4	5	1
Non	16	11	10	4
Compte-tenu du fait de n'avoir proposé qu'un seul type de composition écrite, que pensez-vous de la variété des sujets?				
bonne	13	11	11	4
satisfaisante	7	4	2	1
non satisfaisante	0	0	2	0
Numérotez les critères suivants de 1 à 5, par ordre d'importance décroissante, pour l'évaluation de ce type de composition écrite. (Note: ici on rapporte la <u>moyenne</u> des réponses.)				
structure générale	3.0	2.2	2.7	3.5
présentation logique des arguments	2.9	2.6	2.4	3.5
style, principalement dans la phrase	2.5	2.8	3.1	2.8
technique de la langue: grammaire, usage, mécanique	3.0	3.2	2.8	1.8
choix des mots	3.5	3.9	4.4	3.3

TABLEAU 8.7

MOYENNE DES ESTIMATIONS DU NIVEAU ATTEINT
A LA FIN DU COURS*
(OBJECTIFS COMMUNS AU PROJET NO. II ET PROJET NO. III)

<u>Objectif</u>	<u>13ème année</u>	<u>12ème année</u>	
		<u>niveau avancé</u>	<u>niveau général</u>
A. Langue			
1. Etude des éléments fondamentaux de la langue:			
1.1 orthographe	2.2	2.5	1.9
1.2 vocabulaire	2.5	2.2	1.9
1.3 grammaire	2.0	2.2	1.6
1.4 stylistique (structures de phrases, phraséologie etc.)	2.5	2.5	1.6
2. Travaux écrits:			
2.2 organisation	2.7	2.5	2.0
2.3 rédaction	2.5	2.5	1.9
2.4 analyse et commentaire de textes	2.3	2.4	1.5

*ECHELLE: 0 - niveau inacceptable
1 - niveau acceptable
2 -
3 - niveau très acceptable
4 -
5 - niveau supérieur
6 -
7 - niveau idéal

CHAPTER NINE

ENGLISH AS A SECOND LANGUAGE*

1. INTRODUCTION

The Project II testing of English as a second language was complicated by many of the same factors which affected first language testing in both French and English. The original intention of the project staff had been to test only students in Grade Thirteen, but many of the Francophone schools expressed a desire to have Grade Twelve students included in the test population. Therefore, as in the case of first language testing, the project staff were faced with a rather disparate group of students, who were studying English at several different levels. There were Grade Thirteen students, Grade Twelve students of "English" (the more advanced course), and Grade Twelve students of "Anglais" (the less advanced course). Project III secondary school data were, however, collected only at the Grade Thirteen level.

A further complication of data gathering in this case was that the frame of reference for this testing was somewhat different than for most other subjects. What was of interest was not whether Francophone students of English were sufficiently competent in the language to pursue English studies at the

*(This chapter was written with the assistance of Stephen Ituen, a teacher of both French and English as second languages in Togo, who is presently studying at OISE.)

postsecondary level, but rather whether their command of English was sufficient to enable them to pursue studies in any discipline at an English-language postsecondary institution. The assumption underlying the choice of this frame of reference was that, because of the relative paucity of choice in Ontario of postsecondary courses of study in French, many Francophone students would be required to enrol in English-language institutions.

The effect of this focus on the Project III data-gathering was that all data at the postsecondary level were gathered from teachers of remedial English courses designed for Francophone students. Consequently, these postsecondary data, although of great importance to the Project III report, are of little use in interproject analysis, since they are concerned with those students in need of remedial assistance rather than with the average student entering an English-language institution. In the case of Project II data, the effect was on the wording and distribution at the postsecondary level of the Test Appraisal Inventories which accompanied the test. Although they were distributed to teachers of English at certain colleges and universities, these teachers were asked to respond to the test as a test of language capability for postsecondary study at their institution in general, rather than as a test of capability for further study of the English language.

These two factors--the diversity of the population tested and the focus on general language capability rather than specific curriculum content--resulted in the choice of a test which was not directly tied to the curriculum. No attempt was made, for example, to test literature (which, to judge from Project III data, constituted about 40% of the average Grade Thirteen course). There were severe limitations on the amount of time available for test selection and/or development, and limitations on the amount of time available in the schools on the testing day; consequently it proved impossible to include tests of speaking and listening in the study. This was an unavoidable but regrettable omission.

The only readily available test of vocabulary was pretested in a small number of classrooms, and proved to be much too easy for this population. (This is in itself rather cheering news, since it forms part of a test widely used throughout North America to measure the English-language capability of foreign students entering universities.) The time limitations mentioned earlier did not permit the development of a more suitable instrument.

What was tested were two basic skills--reading comprehension and writing. The Test de compréhension de la langue (anglais), as it finally emerged, had two parts. The first was another section of the test mentioned above, made up of four short reading comprehension passages, each with five related multiple-choice questions. The second presented the student with a somewhat longer passage, argumentative in nature, excerpted from a Canadian magazine, and required him/her to write a 150 to 200 word summary of the passage and then briefly state a personal opinion on the question under discussion. Thus this second part of the test required the student to demonstrate both the ability to read the passage with understanding and the ability to write effectively in two different ways.

2. THE DATA

Table 9.1 contains a list, excerpted from the Project III Course Description Questionnaire, of possible objectives of a Grade Thirteen course in English for Francophone students. Those objectives followed by an asterisk are the ones touched upon in the Project II testing program. As is evident from the table, a great many of these objectives have not been tested, for reasons outlined earlier in this chapter.

It is of some interest, in spite of the intentional lack of curriculum orientation of the test, to examine in rough terms what proportion of the average Grade Thirteen English course is

covered by the test. Project III data indicate that about 38% of total course time is spent on literature, about 9% on other media, and about a further 9% on language study--all areas excluded from testing. Of the approximately 44% of time devoted to language skills, about 11% and 5% are spent on speaking and listening respectively. The areas of reading and writing, which are those tested, are left with about 8% and 20% of class time respectively. It must also be noted that the lists of objectives under these headings include a number of untested skills (e.g., flexibility in reading speed, facility in writing a business letter). It is safe to conclude that the proportion of the average course covered by the test is well under one-quarter.

The structure of the test also makes it very difficult to make direct comparisons between the responses of teachers to the Project III questionnaire and both test results and teacher responses to the Project II Test Appraisal Inventory (described below). Although the topics starred in the list in Table 9.1 are all tested to some extent, they have not in most cases been examined in isolation. For example, student performance on objective 8 is clearly highly relevant to their ability to carry out the writing task presented by the test, but no separate analysis was done of grammatical and structural correctness in these essays. Consequently, although we shall report Project III data related to these objectives where they seem relevant, it must be borne in mind that the fit between the two sets of data is poor.

In addition to the data reported above on percentages of class time allotted to various objectives, the Project III data of importance to this chapter are the assessments by Grade Thirteen teachers of the level of competence of their students in dealing with each objective included in the Project III list and tested by Project II. These assessments were made on the following scale:

0 - no competence

1 - minimal competence

2 - moderate competence

3 - competence in varied situations, some originality

4 - mastery, competence in high level, creative situations

What will be reported here are the means of all the individual assessments.

Project II data are of two kinds. The first is information about student performance on the test. The second consists of responses of teachers to Test Appraisal Inventories, designed to allow secondary and postsecondary teachers to comment both on the test itself and on their own teaching practices and expectations in the areas of English covered by the test.

3. READING COMPREHENSION

3.1 Project III Questionnaire Data

Three of the objectives included in the Project III list under the heading of reading were involved in the reading comprehension part of the test. Grade Thirteen teachers expressed fairly high expectations of their students in these areas. Exit competence in objective I-1, "Comprehend the meaning of a passage", was rated as 2.8, approaching "competence in varied situations, some originality". Ability to distinguish between essential and non-essential information (objective I-2) was rated almost as highly, at 2.6. Slightly lower at 2.2, just above "moderate competence", was objective I-4, "Apply inferential skills".

3.2 Test Appraisal Inventory Data

Table 9.2 reports the answers of secondary and postsecondary teachers to a number of questions on the Test Appraisal Inventories dealing with both the reading comprehension section of the test and more general issues to do with reading comprehension skills. The data are reported separately for Grade Thirteen teachers, Grade Twelve teachers of English, and Grade Twelve teachers of Anglais; the responses of university and community college teachers are combined.

The great proportion of teachers in each of these groups felt that a test of reading comprehension assessed an important component of language achievement. The overall difficulty of the passages in this test was considered to be slightly low by Grade Thirteen teachers, about right by teachers of Grade Twelve English, and slightly high by teachers of Grade Twelve Anglais. Postsecondary teachers felt the passages were about right in difficulty. The passages were generally considered, except at the level of Grade Twelve Anglais, to be representative of the material students should be able to read with comprehension.

Teachers were asked about three particular components of reading comprehension. Postsecondary teachers were unanimous in expecting all incoming students to be able to read a passage for literal understanding. Almost all Grade Thirteen teachers felt that more than three-quarters of their students should have this skill on entry to their courses; two-thirds of Grade Twelve English teachers made the same statement. Grade Twelve Anglais teachers were more pessimistic; only four of ten felt that more than three-quarters of their incoming students should have this ability.

Almost all postsecondary teachers expected incoming students to be able to identify the main idea or purpose of a passage. Fewer teachers at all secondary levels expected this skill than the previous one of most of their incoming students, but

expectations were still fairly high at the Grade Thirteen and Grade Twelve English levels.

Expectations from all four groups of teachers were much lower when it came to drawing inferences and seeing implications. Fewer than half of any of the three secondary groups expected this capability from three-quarters or more of their students. Postsecondary expectations were still fairly high, with 11 of these 15 teachers feeling that more than three-quarters of their incoming students should be able to draw inferences and see implications.

In Grade Thirteen, teachers reported giving the greatest emphasis to reading to identify the main idea or purpose, and to inference and implication. A number of Grade Thirteen teachers gave only light or remedial emphasis to reading for literal understanding (presumably assuming that students should already be competent in this on entry to Grade Thirteen). In Grade Twelve English classes, much more emphasis was given to literal understanding, with identification of the main idea or purpose a close second, and somewhat less attention paid to inference and implication. Grade Twelve Anglais teachers paid most attention to identification of the main idea or purpose and least to reading for literal understanding.

At all but the Grade Twelve Anglais level, teachers felt that other reading skills should have been tested as well. The technical report on the test (Appendix A4 to the Project II report) lists some of the suggestions made. Most teachers felt--many with qualifications (also reported in the above-mentioned appendix)--that the multiple-choice format was a suitable method of measuring those skills that were tested.

Teacher assessments of the difficulty level of each individual passage are reported in detail in Table 9.2. Table 9.3 reports for each of these passages the mean assessment of each passage by each group of teachers, and for each related test item the percentage of Grade Thirteen students and of Grade

Twelve students correctly answering the item. One item has been excluded from the analysis because it was the consensus of the Project II staff that this item offered the student two correct choices.

In each case, Grade Twelve Anglais teachers rated the passages as somewhat more difficult than did any other group. The assessments of the three groups of teachers show some variability, although in three of four cases the Grade Thirteen teachers considered the passage easier than did any other group; in the fourth case, the lowest assessment of difficulty was by the Grade Twelve English teachers. It is a little surprising that except in one case postsecondary teachers considered the passages more difficult than did either Grade Thirteen teachers or Grade Twelve English teachers; this may be attributable to the joint reporting of assessments by university and college teachers, with somewhat lower expectations likely at the college level.

In rough terms, Grade Thirteen teachers felt two passages were somewhat easy, and two about right in difficulty. Grade Thirteen students did very well on the two passages considered easy, with mean success rates of 80% and 81%. However, they did almost equally well on one of the other passages, with a mean success rate of 78%. The remaining passage (number 2) was considerably less well done, with a mean success rate of 64%.

Two passages were considered by both groups of Grade Twelve teachers to be of slightly more than appropriate difficulty. The remaining two were judged to be about right by teachers of Anglais, and somewhat easy by teachers of English. As for Grade Thirteen, performance was relatively good on the two passages judged easier, but almost as good on one of those judged more difficult; students did considerably more poorly on the remaining passage (again number 2). Success rates for Grade Twelve students averaged about 10% lower than those for Grade Thirteen students--not an alarming difference, considering that the Grade Twelve sample contained Anglais students as well as English students.

Most of the reading comprehension items tested literal understanding of the passages. Item 16 asked students to identify the main idea of the passage; students at both levels did well on this item. Four items (numbers 11,12, 17 and 20) required some inference, though not at a high level. Performance on these items was at about the same level as on those demanding only literal understanding.

The ability of students in both Grade Twelve and Grade Thirteen to read with understanding appears, so far as can be judged from these test results, to be acceptably high, and to justify the confidence of their teachers in their skill.

4. WRITING EXERCISE

4.1 Project III Questionnaire Data

A number of the objectives listed in the Project III questionnaire have a bearing on this part of the test. We have already discussed above the data pertaining to objectives in reading. In writing, Grade Thirteen teachers assessed student ability in writing an effective summary of materials read (objective I-5) as 2.9--close to "competence in varied situations, some originality". A similar rating (2.9) was given to the very similar objective 6f. Their ability to demonstrate facility in writing a persuasive or argumentative essay was put slightly lower at 2.5, still better than moderate competence. In the same range were teacher assessments of objectives I-8 (2.5), I-9 (2.6), I-10 (2.8) and I-11 (2.4), which are more specific and limited. Taking all these ratings into account, Grade Thirteen teachers appear to have felt confident that their students were capable in the sort of writing demanded in this test.

4.2 Test Appraisal Inventory Data

Table 9.4 gives the answers of teachers at all four levels of interest to the Test Appraisal Inventory questions on the writing exercise, and on writing in general.

Postsecondary teachers saw an assessment of a sample of a student's writing as essential in evaluating his/her language competence; all groups of secondary teachers considered it important but not essential. Postsecondary expectations that students would be competent in this sort of task on entry to their institutions were much higher than those of secondary teachers that students would have this competence on leaving their classes. Of 16 postsecondary instructors, 15 felt that more than three-quarters of their incoming students should have this skill; even at the Grade Thirteen level, only 13 of 18 teachers felt that more than three-quarters of their students left them with this capability, and only half of the Grade Twelve Anglais teachers made this claim. In view of the high postsecondary expectations, it may be unfortunate that so few secondary teachers reported giving heavy emphasis to developing this competence, with a number saying they gave only light emphasis to it.

In general, teachers at all levels except Grade Twelve Anglais felt that the level of difficulty of the assignment was reasonable, and that the restriction to a single topic was not unfair to the students. When asked to rank in order of importance five possible criteria for evaluating the type of writing asked for in the test, the four groups of teachers were unanimous in placing "diction" in last place. Grade Thirteen teachers and Grade Twelve English teachers were most concerned with "logic, use of evidence" and "organization"; Grade Twelve Anglais teachers gave most weight to "organization" and "grammar, usage and mechanics". All three of these criteria were given about equal weight by postsecondary teachers.

The apparent discrepancy between the expectations reported on the Test Appraisal Inventory and those reported on the Project III questionnaire are at least in part due to the limitation of distribution of the questionnaire to Grade Thirteen teachers; these teachers' responses to the inventory, when considered alone, are more consistent with the questionnaire responses.

It would appear from these data that good performance should be expected from Grade Thirteen students and from Grade Twelve English students; it would be likely that Grade Twelve Anglais students would do more poorly on this task.

4.3 Student Performance

Because of the often-mentioned time limitations of the study, it was not possible to perform an extensive analysis of the writing produced by students. Each test paper was marked independently by three markers; as a check on intermarker reliability, a sample of 20 papers was marked by all six markers employed. Two different marks were assigned to each paper by each marker; these were then averaged over markers to produce a pair of final scores for each paper. The first was a mark for the summary portion of the exercise. A student was given one mark for each main point included in his/her summary to a maximum of ten; marks were deducted for erroneous or unclearly stated points or for variation from the limits imposed on length. The second was a holistic mark on the second part of the exercise, in which the student was to state a personal point of view on the issue raised by the reading passage. A scale from 1 to 10 was employed, and the markers were instructed to use the entire scale; thus no particular mark should be interpreted as a "passing" mark, as each mark has meaning only relative to other marks assigned (as indicating "better" or "worse" performance rather than as a black-and-white judgement of quality).

The only analysis it proved possible to do on these exercises was a comparison of the marks of five different groups of students:

- (a) SSHGD (Grade Thirteen) students planning to proceed to postsecondary education (usually university);
- (b) SSHGD students with other plans for September 1976;
- (c) SSGD (Grade Twelve) students remaining in secondary school;
- (d) SSGD students planning to proceed to postsecondary education (usually community college); and
- (e) SSGD students with other plans for September 1976.

Details of these score distributions are presented in the Project II report; a few general statements will be made here.

On the first, or summary, part of the task, performance by SSHGD students proceeding to postsecondary education was generally better than that of any other group. However, the best of the SSGD students remaining in secondary school proved to receive higher scores than the best of that first group. The SSGD students leaving the educational system performed consistently more poorly than any others.

The situation was somewhat different on the second, or commentary, part. SSHGD students leaving the educational system performed at very close to the same level as those proceeding to postsecondary education. SSGD students remaining in secondary school did considerably more poorly, and those leaving the system or proceeding to postsecondary education much more poorly still.

A very tentative judgement on the basis of these score distributions is that, although students demonstrate the sort of performance pattern in the summary task that is consistent with

the differences in level of their academic expectations, this does not hold for the less routine task presented to them. There are very marked differences in performance in this somewhat less restricted and more creative mode of writing. We repeat our earlier statement that the relative lack of emphasis given to this sort of writing is unfortunate.

5. SUMMARY

In brief, the students tested performed at what appears to be a quite acceptable level in reading. Their ability to write a summary of the longer passage presented to them in the second half appeared to be somewhat less than their reading ability, but performance was consistent and not unacceptable. Great variation was manifested in the commentary portion of this second half, with Grade Twelve performance, particularly among those students probably enrolled in Anglais rather than English, far below the quite adequate standard set by the Grade Thirteen students.

TABLE 9.1

OBJECTIVES LISTED IN COURSE DESCRIPTION QUESTIONNAIRE

I. Language Skills

In reading:

1. Comprehend the meaning of a passage. *
2. Distinguish between essential and non-essential information. *
3. Apply flexibility in the speed of reading appropriate to content and purpose.
4. Apply inferential skills. *

In writing:

5. Write an effective summary of material read. *
6. Demonstrate facility in writing in terms of planning, organization, presentation and editing:
 - a) expository writing
 - b) persuasive or argumentative essay *
 - c) business letter
 - d) personal letter
 - e) report
 - f) summary *
7. Demonstrate facility in imaginative, creative writing:
 - a) prose
 - b) poetry
8. Apply appropriate structure, grammar, and conventions of written English. *

TABLE 9.1 (continued)

9. Use correct spelling. *
 10. Demonstrate effective organization in the writing of a paragraph. *
 11. Present an argument effectively. *
 12. Use effective note-taking techniques.
- In speaking:
13. Present an oral summary of a speech, story or report.
 14. Speak in a style appropriate to subject matter and audience.
 15. Present the following effectively:
 - a) an argument, formal debate
 - b) a story to amuse or entertain
 - c) a formal speech (e.g., in public speaking)
 16. Contribute effectively in a small group discussion
 17. Use language appropriately in a variety of social contexts.

II. Language Study

1. Analyse language in terms of grammar.
2. Apply appropriate terminology in analysis of language.
3. Analyse problems arising from French and English language differences (idioms, imagery, etc.)

TABLE 9.1 (continued)

III. Media Other Than Literature

1. Critically assess
 - a) film
 - b) magazines
 - c) newspapers
 - d) television/radio
 - e) drama
2. Critically assess advertising in terms of:
 - a) language used
 - b) design and presentation (oral and/or written)
3. Critically assess the language of politics (e.g., political speeches, editorials) in terms of:
 - a) language used
 - b) design and presentation
4. Contribute to the presentation of:
 - a) a film
 - b) a videotape or radio presentation
 - c) a play
 - d) an oral interview

IV. Literature

1. Apply a critical vocabulary in the evaluation of the range, nature and quality of a particular work.
2. Analyse literary forms in terms of:
 - a) stylistic techniques (e.g., plot, style, characterization)
 - b) connotation (e.g., word, image, sound)

TABLE 9.2

RESPONSES OF TEACHERS TO QUESTIONS ON
READING COMPREHENSION AND ON PART I OF
THE TEST DE CONNAISSANCE DE LA LANGUE (ANGLAIS)

	<u>Grade 13</u>	<u>Grade 12 Eng.</u>	<u>Grade 12 Ang.</u>	<u>college and university</u>
Do you consider that a test of reading compre- hension, not necessarily the one used, assesses an important component of language achievement at this level?				
Yes	17	23	10	18
No	4	0	1	0
What is your overall assessment of the difficulty level of the four reading comprehen- sion passages?				
too easy	1	2	0	0
somewhat easy	5	2	2	4
about right	13	14	4	7
somewhat difficult	0	4	3	5
too difficult	0	1	1	1
What is your assessment of the difficulty level of the first passage?				
too easy	2	2	0	3
somewhat easy	13	9	2	4
about right	3	10	6	7
somewhat difficult	0	2	2	2
too difficult	0	0	0	2
. . . the second passage?				
too easy	0	1	0	1
somewhat easy	7	4	2	1
about right	4	6	1	4
somewhat difficult	7	8	5	10
too difficult	0	4	2	2

TABLE 9.2 (continued)

	<u>Grade 13</u>	<u>Grade 12 Eng.</u>	<u>Grade 12 Ang.</u>	<u>college and university</u>
. . . the third passage?				
too easy	1	0	0	1
somewhat easy	9	6	1	4
about right	8	8	8	9
somewhat difficult	0	8	1	3
too difficult	0	1	0	1
. . . the fourth passage?				
too easy	0	0	0	1
somewhat easy	4	3	1	0
about right	7	8	2	10
somewhat difficult	7	7	6	6
too difficult	0	5	1	1
Are the four passages representative of the material you would expect students at this level to be able to read with comprehension?				
Yes	15	14	4	12
No	3	8	5	6
How many of the students (<u>entering</u> secondary courses at this level) / (<u>beginning</u> postsecondary studies at your college or university) should have the ability to read a passage for literal understanding?				
100%	8	5	0	16
76%-99%	10	11	4	0
51%-75%	1	6	4	0
26%-50%	0	2	1	0
1%-25%	0	0	1	0
none	0	0	0	0

TABLE 9.2 (continued)

	<u>Grade 13</u>	<u>Grade 12 Eng.</u>	<u>Grade 12 Ang.</u>	<u>college and university</u>
How many of the students (<u>entering</u> secondary courses at this level) / (<u>beginning</u> postsecondary studies at your college or university) should be able to identify the main idea or purpose of a passage?				
100%	1	6	0	12
76%-99%	14	7	3	3
51%-75%	3	7	5	0
26%-50%	1	3	2	1
1%-25%	0	0	0	0
none	0	1	0	0
How many of the students (<u>entering</u> secondary courses at this level) / (<u>beginning</u> postsecondary studies at your college or university) should be able to draw inferences and see implications?				
100%	1	2	0	8
76%-99%	7	6	2	3
51%-75%	10	6	4	2
26%-50%	0	5	4	1
1%-25%	0	4	0	1
none	0	0	0	0
What emphasis is given (in your course) to reading for literal understanding?				
heavy emphasis	4	10	2	--
moderately heavy emphasis	8	6	3	--
light emphasis	4	2	5	--
individual, remedial emphasis only	3	5	0	--
no emphasis	0	0	0	--

TABLE 9.2 (continued)

	<u>Grade 13</u>	<u>Grade 12 Eng.</u>	<u>Grade 12 Ang.</u>	<u>college and university</u>
What emphasis is given (in your course) to reading to identify the main idea or purpose?				
heavy emphasis	8	9	5	--
moderately heavy emphasis	11	8	4	--
light emphasis	0	6	0	--
individual, remedial emphasis only	0	0	1	--
no emphasis	0	0	0	--
What emphasis is given (in your course) to drawing inferences and seeing implications?				
heavy emphasis	7	9	3	--
moderately heavy emphasis	11	5	4	--
light emphasis	1	8	3	--
individual, remedial emphasis only	0	0	0	--
no emphasis	0	1	0	--
Are there important reading skills that have not been tested in the tests used but that should have been?				
Yes	6	9	2	9
No	10	9	6	8
Is the multiple-choice format a reasonable method of assessing at least those three reading comprehension abilities mentioned above?				
Yes	9	12	2	8
Yes (Qualified)	8	8	7	8
No	2	3	0	1

TABLE 9.3
SUMMARY DATA: READING COMPREHENSION PASSAGES

Passage	Mean assessment of difficulty*				Related test items	% correct	
	Grade 13	Grade 12 English	Grade 12 Anglais	college and university		Grade 13	Grade 12
1	2.1	2.5	3.0	2.8	1	90	82
					2	61	56
					3	74	55
					4	84	78
					5	93	81
					Mean	80	70
2	3.0	3.4	3.7	3.6	6	83	63
					7	28	25
					8	80	69
					9	**	**
					10	65	51
					Mean	64	52
3	2.4	2.2	3.0	2.9	11	89	85
					12	74	55
					13	98	97
					14	75	65
					15	67	57
					Mean	81	72

TABLE 9.3 (continued)

Passage	Mean assessment of difficulty*				Related test items	% correct	
	Grade 13	Grade 12 English	Grade 12 Anglais	college and university		Grade 13	Grade 12
4	3.2	3.6	3.7	3.3	16 17 18 19 20 Mean	80 66 70 94 78 78	74 55 59 87 61 67

- *1--too easy
- 2--somewhat easy
- 3--about right
- 4--somewhat difficult
- 5--too difficult

**This item was excluded from the scoring because it was judged to offer the student two correct responses.

TABLE 9.4

RESPONSES OF TEACHERS TO QUESTIONS ON WRITING
AND ON PART II OF THE
TEST DE CONNAISSANCE DE LA LANGUE (ANGLAIS)

	<u>Grade 13</u>	<u>Grade 12 Eng.</u>	<u>Grade 12 Ang.</u>	<u>College and univ.</u>
What importance do you place upon a sample of the student's writing in an evaluation of language competence?				
essential	7	7	2	14
important but not essential	10	14	5	3
of minimal importance and utility	1	2	3	1
neither important nor useful	1	0	0	0
How many students should be able to produce an acceptable piece of writing of this type (on entry to secondary courses at this level)/ (on beginning their post-secondary studies at your college or university)?				
100%	2	3	0	7
76%-99%	6	3	0	8
51%-75%	10	11	5	1
26%-50%	0	2	4	0
1%-25%	0	3	0	0
none	0	1	1	0
How much emphasis do you give in your teaching to the development of student competence in this type of writing?				
Heavy emphasis	3	5	2	--
Moderately heavy emphasis	10	12	3	--
Light emphasis	5	6	4	--
No emphasis	0	0	1	--

TABLE 9.4 (continued)

	<u>Grade 13</u>	<u>Grade 12 Eng.</u>	<u>Grade 12 Ang.</u>	<u>College and univ.</u>
How many of the students who successfully complete courses at (your) level should be able to produce an acceptable piece of writing of this type?				
100%	3	5	2	--
76%-99%	10	12	3	--
51%-75%	5	6	4	--
26%-50%	0	0	1	--
1%-25%	0	0	0	--
none	0	0	0	--
Was the assignment at a reasonable level of difficulty for students (in secondary courses at this level)/(beginning post-secondary studies at your college or university)?				
Yes	16	17	4	16
No	2	6	4	2
Was the restriction to a single topic unfair to students in a test of writing competence?				
Yes	4	6	5	2
No	14	17	4	16

TABLE 9.4 (continued)

	<u>Grade 13</u>	<u>Grade 12 Eng.</u>	<u>Grade 12 Ang.</u>	<u>College and univ.</u>
<p>Rank, in order of importance, the following criteria for evaluating this type of writing (1--High, 5--Low). (Note: What are reported here are the <u>mean</u> rankings.)</p>				
Organization	2.6	2.2	1.8	2.4
Logic, use of evidence	2.2	2.1	3.0	2.3
Style (chiefly the sentence)	3.0	3.2	3.2	3.3
Grammar, usage, mechanics	2.8	2.8	2.2	2.2
Diction	3.8	4.1	3.7	3.4

